

SEARCH REQUEST FORM**Scientific and Technical Information Center**

Requester's Full Name: _____ Examiner #: _____ Date: _____
 Art Unit: _____ Phone Number 30 _____ Serial Number: _____
 Mail Box and Bldg/Room Location: _____ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____

Inventors (please provide full names): _____

Earliest Priority Filing Date: _____

**For Sequence Searches Only* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

STAFF USE ONLY**Type of Search****Vendors and cost where applicable**

Searcher: Baker, E 4994 NA Sequence (#) _____ STN _____
 Searcher Phone #: _____ AA Sequence (#) _____ Dialog _____
 Searcher Location: _____ Structure (#) _____ Questel/Orbit _____
 Date Searcher Picked Up: _____ Bibliographic _____ Dr. Link _____
 Date Completed: 01-23-02 Litigation _____ Lexis/Nexis _____
 Searcher Prep & Review Time: 3 Fulltext _____ Sequence Systems _____
 Clerical Prep Time: _____ Patent Family _____ WWW/Internet _____
 Online Time: 20 Other _____ Other (specify) CGN

piegler
09/18/86034
Seq. ID 2 w/ Interf

GenCode version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2002, 15:18:25 ; Search time 96.29 Seconds
(without alignments) 801.627 Million cell updates/sec

Title: US-09-866-034-2
Perfect score: 1399
Sequence: 1 MHRKAGLLGLCARANNSVRMA.....DASYITGETVVVGGTSPRL 278

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 3148936 seqs, 277657034 residues
Total number of hits satisfying chosen parameters: 3148936

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_Main: *
1: /cgn2_6/ptodata/2/paa/US06_COMB.pcp.*
2: /cgn2_6/ptodata/2/paa/US06_COMB.pcp.*
3: /cgn2_6/ptodata/2/paa/US07_COMB.pcp.*
4: /cgn2_6/ptodata/2/paa/US08_COMB.pcp.*
5: /cgn2_6/ptodata/2/paa/US081_COMB.pcp.*
6: /cgn2_6/ptodata/2/paa/US082_COMB.pcp.*
7: /cgn2_6/ptodata/2/paa/US083_COMB.pcp.*
8: /cgn2_6/ptodata/2/paa/US084_COMB.pcp.*
9: /cgn2_6/ptodata/2/paa/US085_COMB.pcp.*
10: /cgn2_6/ptodata/2/paa/US086_COMB.pcp.*
11: /cgn2_6/ptodata/2/paa/US087_COMB.pcp.*
12: /cgn2_6/ptodata/2/paa/US088_COMB.pcp.*
13: /cgn2_6/ptodata/2/paa/US089_COMB.pcp.*
14: /cgn2_6/ptodata/2/paa/US090_COMB.pcp.*
15: /cgn2_6/ptodata/2/paa/US091_COMB.pcp.*
16: /cgn2_6/ptodata/2/paa/US092_COMB.pcp.*
17: /cgn2_6/ptodata/2/paa/US093_COMB.pcp.*
18: /cgn2_6/ptodata/2/paa/US094_COMB.pcp.*
19: /cgn2_6/ptodata/2/paa/US095_COMB.pcp.*
20: /cgn2_6/ptodata/2/paa/US096_COMB.pcp.*
21: /cgn2_6/ptodata/2/paa/US097_COMB.pcp.*
22: /cgn2_6/ptodata/2/paa/US098_COMB.pcp.*
23: /cgn2_6/ptodata/2/paa/US099_COMB.pcp.*
24: /cgn2_6/ptodata/2/paa/US60_COMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1399	100.0	278	22	US-09-866-034-2
2	1399	100.0	278	23	US-09-927-796-60
3	1353	96.7	518	1	PCT-US01-14827-15496
4	1327	94.9	278	1	PCT-US99-16164-1
5	1327	94.9	278	45	US-09-116-750-1
6	1229.5	87.9	477	1	PCT-US01-08656-7140
7	1229.5	87.9	477	1	PCT-US01-14827-9680
8	1229.5	87.9	477	1	PCT-US01-14827-15494
9	1210	86.5	268	1	PCT-US01-08656-10624

10	1196	85.5	244	24	US-60-323-349-175	Sequence 175, Appl
11	1196	85.5	333	1	PCT-US01-08656-7142	Sequence 7142, Ap
12	1161	83.0	440	1	PCT-US01-14827-9686	Sequence 9686, Ap
13	906	64.8	193	1	PCT-US01-01239-1013	Sequence 1013, Ap
14	906	64.8	193	1	PCT-US01-01308-456	Sequence 456, App
15	906	64.8	193	1	PCT-US01-01349-673	Sequence 673, App
16	906	64.8	193	21	US-09-764-849-456	Sequence 456, App
17	906	64.8	193	21	US-09-764-853-673	Sequence 673, App
18	906	64.8	193	21	US-09-764-902-1013	Sequence 1013, Ap
19	824.5	58.9	280	1	PCT-US99-16164-5	Sequence 5, Appli
20	824.5	58.9	280	15	US-09-116-750-5	Sequence 5, Appli
21	824.5	58.9	280	24	US-60-277-380-17	Sequence 17, Appl
22	809.5	57.9	205	1	PCT-US01-14827-9684	Sequence 9684, Ap
23	797	57.0	277	24	US-60-245-241-126	Sequence 126, App
24	677	48.4	136	24	US-60-245-241-127	Sequence 127, App
25	664	47.5	317	24	US-60-167-217-22643	Sequence 22643, A
26	664	47.5	317	24	US-60-173-464-18416	Sequence 18416, A
27	664	47.5	317	24	US-60-191-637-22451	Sequence 22451, A
28	664	47.5	317	24	US-60-191-681-17735	Sequence 17735, A
29	649.5	46.4	183	1	PCT-US00-05881-530	Sequence 530, App
30	649.5	46.4	183	23	US-09-925-238-530	Sequence 530, App
31	620.5	44.4	276	24	US-60-324-109-26991	Sequence 26991, A
32	618.5	44.2	253	24	US-60-324-109-25978	Sequence 25978, A
33	616.5	44.1	265	24	US-60-324-109-16907	Sequence 16907, A
34	588.5	42.1	255	19	US-09-595-329A-1555	Sequence 1555, Ap
35	588.5	42.1	277	19	US-09-595-329A-1554	Sequence 1554, Ap
36	586.5	41.9	263	24	US-60-324-109-22645	Sequence 22645, A
37	468	33.5	221	21	US-09-760-475-1902	Sequence 1902, Ap
38	428	30.6	103	1	PCT-US00-26524B-5344	Sequence 5344, Ap
39	409	29.2	256	16	US-09-252-991A-30474	Sequence 30474, A
40	404	28.9	248	17	US-09-305-390-5	Sequence 5, Appli
41	404	28.9	248	17	US-09-305-390-5	Sequence 5, Appli
42	395.5	28.3	249	23	US-09-902-540-14561	Sequence 14561, A
43	373	26.7	249	15	US-09-134-001C-4825	Sequence 4825, Ap
44	373	26.7	249	18	US-09-450-969-5675	Sequence 5675, Ap
45	373	26.7	287	24	US-60-324-109-29150	Sequence 29150, A

ALIGNMENTS

RESULT 1
US-09-866-034-2
; GENERAL INFORMATION:
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2930RIC1
; CURRENT APPLICATION NUMBER: US/09/866,034
; CURRENT FILING DATE: 2001-05-25
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 38
; SEQ ID NO 2
; LENGTH: 278
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-866-034-2

Query Match 100.0%; Score 1399; DB 22; Length 278;

1	PRIOR APPLICATION NUMBER: 60/097977
2	PRIOR FILING DATE: 1998-08-26
3	PRIOR APPLICATION NUMBER: 60/103396
4	PRIOR FILING DATE: 1998-10-07
5	PRIOR APPLICATION NUMBER: 60/108867
6	PRIOR FILING DATE: 1998-11-17
7	PRIOR APPLICATION NUMBER: 60/112851
8	PRIOR FILING DATE: 1998-12-16
9	PRIOR APPLICATION NUMBER: 60/119965
10	PRIOR FILING DATE: 1999-02-12
11	PRIOR APPLICATION NUMBER: 60/123972
12	PRIOR FILING DATE: 1999-03-11
13	PRIOR APPLICATION NUMBER: 60/133459
14	PRIOR FILING DATE: 1999-05-11
15	PRIOR APPLICATION NUMBER: 60/140650
16	PRIOR FILING DATE: 1999-06-22
17	PRIOR APPLICATION NUMBER: 60/140653
18	PRIOR FILING DATE: 1999-06-22
19	PRIOR APPLICATION NUMBER: 60/144758
20	PRIOR FILING DATE: 1999-07-20
21	PRIOR APPLICATION NUMBER: 60/145698
22	PRIOR FILING DATE: 1999-07-26
23	PRIOR APPLICATION NUMBER: 60/146222
24	PRIOR FILING DATE: 1999-07-28
25	PRIOR APPLICATION NUMBER: 60/149395
26	PRIOR FILING DATE: 1999-08-17
27	PRIOR APPLICATION NUMBER: 60/151689
28	PRIOR FILING DATE: 1999-08-31
29	PRIOR APPLICATION NUMBER: 08/625328
30	PRIOR FILING DATE: 1996-04-01
31	PRIOR APPLICATION NUMBER: 08/710802
32	PRIOR FILING DATE: 1996-09-23
33	PRIOR APPLICATION NUMBER: 08/800699
34	PRIOR FILING DATE: 1997-02-14
35	PRIOR APPLICATION NUMBER: 08/828683
36	PRIOR FILING DATE: 1997-03-31
37	PRIOR APPLICATION NUMBER: 08/829270
38	PRIOR FILING DATE: 1997-03-31
39	PRIOR APPLICATION NUMBER: 08/928069
40	PRIOR FILING DATE: 1997-09-11
41	PRIOR APPLICATION NUMBER: 08/934494
42	PRIOR FILING DATE: 1997-09-19
43	PRIOR APPLICATION NUMBER: 09/143068
44	PRIOR FILING DATE: 1998-08-28
45	PRIOR APPLICATION NUMBER: 09/143707
46	PRIOR FILING DATE: 1998-08-28
47	PRIOR APPLICATION NUMBER: 09/151889
48	PRIOR FILING DATE: 1998-09-11
49	PRIOR APPLICATION NUMBER: 09/169104
50	PRIOR FILING DATE: 1998-10-09
51	PRIOR APPLICATION NUMBER: 09/202089
52	PRIOR FILING DATE: 1998-12-08
53	PRIOR APPLICATION NUMBER: 09/254311
54	PRIOR FILING DATE: 1999-03-03
55	PRIOR APPLICATION NUMBER: 09/304003
56	PRIOR FILING DATE: 1999-04-30
57	PRIOR APPLICATION NUMBER: 09/380137
58	PRIOR FILING DATE: 1999-08-25
59	PRIOR APPLICATION NUMBER: 09/380138
60	PRIOR FILING DATE: 1999-08-25
61	PRIOR APPLICATION NUMBER: 09/380139
62	PRIOR FILING DATE: 1999-08-25
63	PRIOR APPLICATION NUMBER: 09/403297
64	PRIOR FILING DATE: 1999-10-18
65	PRIOR APPLICATION NUMBER: 09/423844
66	PRIOR FILING DATE: 1999-11-12
67	PRIOR APPLICATION NUMBER: 09/511133
68	PRIOR FILING DATE: 2000-02-23
69	PRIOR APPLICATION NUMBER: 09/511631
70	PRIOR FILING DATE: 2000-02-23
71	PRIOR APPLICATION NUMBER: 09/664610
72	PRIOR FILING DATE: 2000-09-18
73	PRIOR APPLICATION NUMBER: 09/665350

```

: PRIOR APPLICATION NUMBER: PCT/US01/21066
: PRIOR FILING DATE: 2001-06-29
: PRIOR APPLICATION NUMBER: PCT/US01/21735
: PRIOR FILING DATE: 2001-07-09
: NUMBER OF SEQ ID NOS: 258
: SEQ ID NO 60
: LENGTH: 278

Query Match      100.0%; Score 1399; DB:23; Length 278;
Best Local Similarity 100.0%; Pred. No. 9.6e-138;
Matches 278; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHKAGLLGLCARAWNSVRWASSGMDTRDPLANKVALVTASTDGIQIGFAIARRLAQDGAHV 60
Db 1 MHKAGLLGLCARAWNSVRWASSGMDTRDPLANKVALVTASTDGIQIGFAIARRLAQDGAHV 60

QY 61 VSSRKQNVDOAVATLQGEGLSVTGTVCVHGKAEDRERLVATAVKLHGIDILVSNAAVN 120
Db 61 VSSRKQNVDOAVATLQGEGLSVTGTVCVHGKAEDRERLVATAVKLHGIDILVSNAAVN 120

QY 121 PFEGSIMDVTEEWBDKTLDINVKAPALMTKAVVPEMEKRGGSVVIVSSIAAFSPSPGFS 180
Db 121 PFEGSIMDVTEEWBDKTLDINVKAPALMTKAVVPEMEKRGGSVVIVSSIAAFSPSPGFS 180

QY 181 PYNVSKTALLGLTKTALIELAPNRINRVNCLAPGLIKTFSRMLWMDKEEESMKETLRLR 240
Db 181 PYNVSKTALLGLTKTALIELAPNRINRVNCLAPGLIKTFSRMLWMDKEEESMKETLRLR 240

QY 241 RLGEPEDCAGIVSFLCSEDASYITGETVVVGGGTPSRL 278
Db 241 RLGEPEDCAGIVSFLCSEDASYITGETVVVGGGTPSRL 278

RESULT 3
PCT-US01-14827-15496
: Sequence 15496, Application PC/TUS0114827
: GENERAL INFORMATION:
: APPLICANT: Hyseq, Inc
: TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
: FILE REFERENCE: 21272-104
: CURRENT APPLICATION NUMBER: PCT/US01/14827
: CURRENT FILING DATE: 2001-05-16
: PRIOR APPLICATION NUMBER: 09/577,408
: PRIOR FILING DATE: 2000-05-18
: NUMBER OF SEQ ID NOS: 16102
: SOFTWARE: Custom
: SEQ ID NO 15496
: LENGTH: 518
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: DOMAIN
: LOCATION: (35)...(53)
: OTHER INFORMATION: GLUCOSE/RIBITOL DEHYDROGENASE FAMILY SIGNATURE domain
: OTHER INFORMATION: identified by eMATRIX, accession number PR00081A, p-value=6.22e-72
: OTHER INFORMATION: 13, raw score of 10.53
: NAME/KEY: DOMAIN
: LOCATION: (34)...(220)
: OTHER INFORMATION: short chain dehydrogenase domain identified by Pfam,
: OTHER INFORMATION: accession name adh_short, E-value=2.7e-60, Pfam score of 213.7
PCT-US01-14827-15496

Query Match      96.7%; Score 1353; DB 1; Length 518;
Best Local Similarity 97.1%; Pred. No. 1.7e-132;
Matches 271; Conservative 0; Mismatches 2; Indels 6; Gaps 1;

QY 1 MHKAGLLGLCARAWNSVRWASSGMDTRDPLANKVALVTASTDGIQIGFAIARRLAQDGAHV 60
Db 2 MHKAGLLGLCARAWNSVRWASSGMDTRDPLANKVALVTASTDGIQIGFAIARRLAQDGAHV 61

QY 61 VSSRKQNVDOAVATLQGEGLSVTGTVCVHGKAEDRERLVATAVKLHGIDILVSNAAVN 120

```


Db 62 VSSRKQNVDOAVATLQEGSLVTVGVCHVGAEDRRLVATAVKLHGGIDILVSNAAVN 121
Qy 121 PFFGSDMDVTEVWDKTLIDINVKAPALMTKAVVPEMEKRGSGVVISSIAAFSPSPGFS 180
Db 122 PFFGSDMDVTEVWDKTLIDINVKAPALMTKAVVPEMEKRGSGVVISSIAAFSPSPGFS 181
Qy 181 PYNVSKTALLGLTKTLAIELAPRNRVNCAPGLIKTSFRLMWMDEKESMKETLRIR 240
Db 182 PYNVSKTVLLGLTKTLAIELAPRNRVNCAPGLIKTSFRLMWMDEKESMKETLRIR 241
Qy 241 RLGPEDCAGIVSFLCSEDASYITGETVTV-----VGGG 273
Db 242 RLGPEDCAGIVSFLCSEDASYITGETVTVNLSVMTGGG 280

RESULT 4

PCT-US99-16164-1
; Sequence 1, Application PC/TUS9916164
; GENERAL INFORMATION:
; APPLICANT: INCYTE PHARMACEUTICALS, INC.
; APPLICANT: BANDMAN, Olga
; APPLICANT: TANG, Y. Tom
; APPLICANT: CORLEY, Neil C.
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: BAUGHN, Mariah R.
; TITLE OF INVENTION: SCAD-RELATED MOLECULES
; FILE REFERENCE: PF-0559 PCT
; CURRENT APPLICATION NUMBER: PCT/US99/16164
; PRIOR FILING DATE: 1999-07-16
; EARLIER APPLICATION NUMBER: 09/116,750; Unassigned
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 278
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte Clone No: 1240869
; PUBLICATION INFORMATION:
PCT-US99-16164-1

Query Match 94.9%; Score 1327; DB 1; Length 278;
Best Local Similarity 96.4%; Pred. No. 3.5e-130;
Matches 268; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 MHKAGLLGLCARAWNSVRMASSGTRRDPLANKVALVYASTDGGIGFAIARRLAQDRAHV 60
Db 1 MHMARLLGLCAWARKSVRMASSTRDRDPLTNKVALVYASTDGGIGFAIARRLAQDRAHV 60
Qy 61 VSSRKQNVDOAVATLQEGSLVTVGVCHVGAEDRRLVATAVKLHGGIDILVSNAAVN 120
Db 61 VSSRKQNVDOAVATLQEGSLVTVGVCHVGAEDRRLVATAVKLHGGIDILVSNAAVN 120
Qy 121 PFFGSDMDVTEVWDKTLIDINVKAPALMTKAVVPEMEKRGSGVVISSIAAFSPSPGFS 180
Db 121 PFFGSDMDVTEVWDKTLIDINVKAPALMTKAVVPEMEKRGSGVVISSIAAFSPSPGFS 180
Qy 181 PYNVSKTALLGLTKTLAIELAPRNRVNCAPGLIKTSFRLMWMDEKESMKETLRIR 240
Db 181 PYNVSKTALLGLTNNTLAIELAPRNRVNCAPGLIKTSFRLMWMDEKESMKETLRIR 240
Qy 241 RLGPEDCAGIVSFLCSEDASYITGETVTVVGGGTPSRL 278
Db 241 RLGPEDCAGIVSFLCSEDASYITGETVTVVGGGTPSRL 278

RESULT 5

US-09-116-750-1
; Sequence 1, Application US/09116750
; GENERAL INFORMATION:

; APPLICANT: Bandman, Olga
; APPLICANT: Tang, Y. Tom
; APPLICANT: Corley, Neil C.
; APPLICANT: Azimzai, Yalda
; APPLICANT: Baughn, Mariah R.
; TITLE OF INVENTION: SCAD-RELATED MOLECULES
; FILE REFERENCE: PF-0559 US
; CURRENT APPLICATION NUMBER: US/09/116,750
; CURRENT FILING DATE: 1998-07-16
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 278
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
; FEATURE:
; OTHER INFORMATION: 1240869
US-09-116-750-1

Query Match 94.9%; Score 1327; DB 15; Length 278;
Best Local Similarity 96.4%; Pred. No. 3.5e-130;
Matches 268; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 MHKAGLLGLCARAWNSVRMASSGTRRDPLANKVALVYASTDGGIGFAIARRLAQDRAHV 60
Db 1 MHMARLLGLCAWARKSVRMASSTRDRDPLTNKVALVYASTDGGIGFAIARRLAQDRAHV 60
Qy 61 VSSRKQNVDOAVATLQEGSLVTVGVCHVGAEDRRLVATAVKLHGGIDILVSNAAVN 120
Db 61 VSSRKQNVDOAVATLQEGSLVTVGVCHVGAEDRRLVATAVKLHGGIDILVSNAAVN 120
Qy 121 PFFGSDMDVTEVWDKTLIDINVKAPALMTKAVVPEMEKRGSGVVISSIAAFSPSPGFS 180
Db 121 PFFGSDMDVTEVWDKTLIDINVKAPALMTKAVVPEMEKRGSGVVISSIAAFSPSPGFS 180
Qy 181 PYNVSKTALLGLTKTLAIELAPRNRVNCAPGLIKTSFRLMWMDEKESMKETLRIR 240
Db 181 PYNVSKTALLGLTNNTLAIELAPRNRVNCAPGLIKTSFRLMWMDEKESMKETLRIR 240
Qy 241 RLGPEDCAGIVSFLCSEDASYITGETVTVVGGGTPSRL 278
Db 241 RLGPEDCAGIVSFLCSEDASYITGETVTVVGGGTPSRL 278

RESULT 6

PCT-US01-08656-7140
; Sequence 7140, Application PC/TUS0108656
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-066
; CURRENT APPLICATION NUMBER: PCT/US01/08656
; CURRENT FILING DATE: 2001-04-16
; PRIOR FILING DATE: 09/522,929
; PRIOR APPLICATION NUMBER: 2000-04-18
; PRIOR FILING DATE: 09/770,160
; NUMBER OF SEQ ID NOS: 10994
; SOFTWARE: Custom
; SEQ ID NO 7140
; LENGTH: 477
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (34)...(52)
; OTHER INFORMATION: GLUCOSE/RIBITOL DEHYDROGENASE FAMILY SIGNATURE domain
; OTHER INFORMATION: identified by eMATRIX, accession number PR00081A, p-value=6.22
; OTHER INFORMATION: 13, raw score of 10.53
; NAME/KEY: DOMAIN
; LOCATION: (33)...(218)
; OTHER INFORMATION: short chain dehydrogenase domain identified by Pfam,

OTHER INFORMATION: accession name adh_short, E-value=1.5e-60, Pfam score of 214.6
PCT-US01-08656-7140

Query Match 87.9%; Score 1229.5; DB 1; Length 477;
Best Local Similarity 90.0%; Pred. No. 1.4e-119;
Matches 251; Conservative 0; Mismatches 3; Indels 25; Gaps 2;

Qy 1 MHKAGLLGLCARAWNSVRMASSGMRTRDPLANKVALVTASTDGGIGFATARRLAQDGAHVY 60
Db 1 MHKAGLLGLCARAWNSVRMASSGMRTRDPLANKVALVTASTDGGIGFATARRLAQDGAHVY 60
Qy 61 VSSRQQNVQDAVATLOGEGLSVTGVCHVGAEDRERLAVATVAVKLHGIDILVSNAAVN 120
Db 61 VSSRQQNVQDAVATLOGEGLSVTGVCHVGAEDRERLAVATVAVKLHGIDILVSNAAVN 120
Qy 121 PFFGSIMDVTEVWMDKTLDINVKAPALMTKAVVPEMERKGGSVVIVSSIAAFSPSPGFS 180
Db 121 PFFGSIMDVTEVWMDKTLDINVKAPALMTKAVVPEMERKGGSVVIVSSIAAFSPSPGFS 180
Qy 181 PYNVSKTALLGLTKTALTELAPRNIRVNCPLAPGLIKTSFSRM-----VGGG 273
Db 181 PYNVSKTALLGLTNLTALTELAPRNIRVNCPLAPGLIKTSFSRM-----VGGG 273
Qy 241 RLGEPCDAGIVSFLCSEDASYITGETVTVVNLNLSVMTGGG 260
Db 241 RLGEPCDAGIVSFLCSEDASYITGETVTVVNLNLSVMTGGG 260

RESULT 7

PCT-US01-14827-9680
Sequence 9680, Application PC/TUS0114827
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 21272-104
CURRENT APPLICATION NUMBER: PCT/US01/14827
CURRENT FILING DATE: 2001-05-16
PRIOR APPLICATION NUMBER: 09/577,408
PRIOR FILING DATE: 2000-05-18
NUMBER OF SEQ ID NOS: 16102
SOFTWARE: Custom
SEQ ID NO 9680
LENGTH: 477
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: DOMAIN
LOCATION: (34)..(52)
OTHER INFORMATION: GLUCOSE/RIBITOL DEHYDROGENASE FAMILY SIGNATURE domain
OTHER INFORMATION: Identified by eMATRIX, accession number PR00081A, p-value=6.226e-13, raw score of 10.53
NAME/KEY: DOMAIN
LOCATION: (33)..(218)
OTHER INFORMATION: short chain dehydrogenase domain identified by Pfam.
OTHER INFORMATION: accession name adh_short, E-value=1.5e-60, Pfam score of 214.6
PCT-US01-14827-9680

Query Match 87.9%; Score 1229.5; DB 1; Length 477;
Best Local Similarity 90.0%; Pred. No. 1.4e-119;
Matches 251; Conservative 0; Mismatches 3; Indels 25; Gaps 2;

Qy 1 MHKAGLLGLCARAWNSVRMASSGMRTRDPLANKVALVTASTDGGIGFATARRLAQDGAHVY 60
Db 1 MHKAGLLGLCARAWNSVRMASSGMRTRDPLANKVALVTASTDGGIGFATARRLAQDGAHVY 60
Qy 61 VSSRQQNVQDAVATLOGEGLSVTGVCHVGAEDRERLAVATVAVKLHGIDILVSNAAVN 120
Db 61 VSSRQQNVQDAVATLOGEGLSVTGVCHVGAEDRERLAVATVAVKLHGIDILVSNAAVN 120
Qy 121 PFFGSIMDVTEVWMDKTLDINVKAPALMTKAVVPEMERKGGSVVIVSSIAAFSPSPGFS 180
Db 121 PFFGSIMDVTEVWMDKTLDINVKAPALMTKAVVPEMERKGGSVVIVSSIAAFSPSPGFS 180

Db 121 PFFGSIMDVTEVWMDKTLDINVKAPALMTKAVVPEMERKGGSVVIVSSIAAFSPSPGFS 180
Qy 181 PYNVSKTALLGLTKTALTELAPRNIRVNCPLAPGLIKTSFSRM-----VGGG 273
Db 181 PYNVSKTALLGLTNLTALTELAPRNIRVNCPLAPGLIKTSFSRM-----VGGG 273
Qy 241 RLGEPCDAGIVSFLCSEDASYITGETVTVVNLNLSVMTGGG 260
Db 241 RLGEPCDAGIVSFLCSEDASYITGETVTVVNLNLSVMTGGG 260

RESULT 8

PCT-US01-14827-15494
Sequence 15494, Application PC/TUS0114827
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 21272-104
CURRENT APPLICATION NUMBER: PCT/US01/14827
CURRENT FILING DATE: 2001-05-16
PRIOR APPLICATION NUMBER: 09/577,408
PRIOR FILING DATE: 2000-05-18
NUMBER OF SEQ ID NOS: 16102
SOFTWARE: Custom
SEQ ID NO 15494
LENGTH: 477
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: DOMAIN
LOCATION: (34)..(52)
OTHER INFORMATION: GLUCOSE/RIBITOL DEHYDROGENASE FAMILY SIGNATURE domain
OTHER INFORMATION: Identified by eMATRIX, accession number PR00081A, p-value=6.22e-13, raw score of 10.53
NAME/KEY: DOMAIN
LOCATION: (33)..(218)
OTHER INFORMATION: short chain dehydrogenase domain identified by Pfam.
OTHER INFORMATION: accession name adh_short, E-value=1.5e-60, Pfam score of 214.6
PCT-US01-14827-15494

Query Match 87.9%; Score 1229.5; DB 1; Length 477;
Best Local Similarity 90.0%; Pred. No. 1.4e-119;
Matches 251; Conservative 0; Mismatches 3; Indels 25; Gaps 2;

Qy 1 MHKAGLLGLCARAWNSVRMASSGMRTRDPLANKVALVTASTDGGIGFATARRLAQDGAHVY 60
Db 1 MHKAGLLGLCARAWNSVRMASSGMRTRDPLANKVALVTASTDGGIGFATARRLAQDGAHVY 60
Qy 61 VSSRQQNVQDAVATLOGEGLSVTGVCHVGAEDRERLAVATVAVKLHGIDILVSNAAVN 120
Db 61 VSSRQQNVQDAVATLOGEGLSVTGVCHVGAEDRERLAVATVAVKLHGIDILVSNAAVN 120
Qy 121 PFFGSIMDVTEVWMDKTLDINVKAPALMTKAVVPEMERKGGSVVIVSSIAAFSPSPGFS 180
Db 121 PFFGSIMDVTEVWMDKTLDINVKAPALMTKAVVPEMERKGGSVVIVSSIAAFSPSPGFS 180
Qy 181 PYNVSKTALLGLTKTALTELAPRNIRVNCPLAPGLIKTSFSRM-----VGGG 273
Db 181 PYNVSKTALLGLTNLTALTELAPRNIRVNCPLAPGLIKTSFSRM-----VGGG 273
Qy 241 RLGEPCDAGIVSFLCSEDASYITGETVTVVNLNLSVMTGGG 260
Db 241 RLGEPCDAGIVSFLCSEDASYITGETVTVVNLNLSVMTGGG 260

RESULT 9

PCT-US01-08656-10624
Sequence 10624, Application PC/TUS0108656
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc
TITLE OF INVENTION: NOVEL MACROPHAGE NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 21272-066

```

CURRENT APPLICATION NUMBER: PCT/US01/08656
CURRENT FILING DATE: 2001-04-16
PRIOR APPLICATION NUMBER: 09/522,929
PRIOR FILING DATE: 2000-04-16
PRIOR APPLICATION NUMBER: 09/770,160
PRIOR FILING DATE: 2001-01-26
NUMBER OF SEQ ID NOS: 10994
SOFTWARE: Custom
SEQ ID NO 10624
LENGTH: 268
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: DOMAIN
LOCATION: (151)...(189)
OTHER INFORMATION: Short-chain dehydrogenases/reductases family proteins domain
OTHER INFORMATION: identified by eMATRIX, accession number BL00061B, p-value=2.500e-
OTHER INFORMATION: 16, raw score of 25.79
NAME/KEY: DOMAIN
LOCATION: (21)...(216)
OTHER INFORMATION: short chain dehydrogenase domain identified by Pfam,
OTHER INFORMATION: accession name adh_short, E-value=6.6e-27, Pfam score of 102.8
NAME/KEY: misc_feature
LOCATION: (1)...(268)
OTHER INFORMATION: Xaa = X or * as defined in Table 2
PCT-US01-08656-10624

Query Match      86.5%; Score 1210; DB1; Length 268;
Best Local Similarity 92.9%; Pred. No. 6.5e-118;
Matches 249; Conservative 1; Mismatches 16; Indels 2; Gaps 2;

QY 13 AINSVRMASSGTRDRPLANKVALVTASTDGIQFAIARRLAQDGAHVVSRRKQONVDQA 72
Db 1 ARKSVRMASSRTRDRPLTNKVALVTASTDGIQFAIARRLAQDGAHVVSRRKQONVDQA 60
QY 73 VATLQEGSLVTGTVCHVKGAEADR-ERLVATAVKLHGIDILVSNAAVNPFGSIMDYTE 131
Db 61 VATLQEGSLVTGTVCHVKGAEADRGAAWPPAVKLHGIDILVSNAAVNPFGSIMDYTE 120
QY 132 EVMDKTLIDINVKAP-ALMTKAVVPEKRRGGSVVIVSSIAAFSPGFSFYNYSKALL 190
Db 121 EVMDKTLIDINVKPKPMXTRKAVPEMEKRRGGSVVIVSSIAAFSPGFSFYNYSKALL 180
QY 191 GLTKLTAIELAPRINRVNCLAPGLIKTSFSRMLWMDKEESMKETLIRRLGPEPCAG 250
Db 181 GLAQLTLELAPRINRVNCLAPGLIKTSFSRMLWMDKEESMKETLIRRLGPEPCAG 240
QY 251 IVSFLCSEDASYITGETVWVGGGTPSRL 278
Db 241 IVSFLCSEDASYITGETVWVGGGTPSRL 268

RESULT 10
US-60-323-349-175
Sequence 175, Application US/60323349
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Asundi, Vinod
APPLICANT: Goodrich, Ryle W
APPLICANT: Ren, Felyan
APPLICANT: Zhang, Jie
APPLICANT: Zhao, Qing A.
APPLICANT: Wang, Jian-Rui
APPLICANT: Ghosh, Malabika
APPLICANT: Xue, Aidong J.
APPLICANT: Wehrman, Tom
APPLICANT: Weng, Gezhi
APPLICANT: Zhou, Ping
APPLICANT: Drmanac, Radoje T
TITLE OF INVENTION: Novel Nucleic Acids and Secreted Polypeptides
FILE REFERENCE: 808
CURRENT APPLICATION NUMBER: US/60/323,349

```

```

CURRENT FILING DATE: 2001-09-19
PRIOR APPLICATION NUMBER: US 09/488,725
PRIOR FILING DATE: 2000-01-21
PRIOR APPLICATION NUMBER: US 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: PCT/US00/35017
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: US 09/491,404
PRIOR FILING DATE: 2000-01-25
PRIOR APPLICATION NUMBER: PCT/US01/02623
PRIOR FILING DATE: 2001-01-25
PRIOR APPLICATION NUMBER: US 09/496,914
PRIOR FILING DATE: 2000-02-03
PRIOR APPLICATION NUMBER: US 09/560,875
PRIOR FILING DATE: 2000-04-27
PRIOR APPLICATION NUMBER: PCT/US01/03800
PRIOR FILING DATE: 2001-02-05
PRIOR APPLICATION NUMBER: US 09/515,126
PRIOR FILING DATE: 2000-02-28
PRIOR APPLICATION NUMBER: US 09/577,409
PRIOR FILING DATE: 2000-05-18
PRIOR APPLICATION NUMBER: PCT/US01/04927
PRIOR FILING DATE: 2001-02-26
PRIOR APPLICATION NUMBER: US 09/519,705
PRIOR FILING DATE: 2000-03-07
PRIOR APPLICATION NUMBER: US 09/574,454
PRIOR FILING DATE: 2000-05-19
PRIOR APPLICATION NUMBER: PCT/US01/04941
PRIOR FILING DATE: 2001-03-05
PRIOR APPLICATION NUMBER: US 09/540,217
PRIOR FILING DATE: 2000-03-31
PRIOR APPLICATION NUMBER: US 09/649,167
PRIOR FILING DATE: 2000-08-23
PRIOR APPLICATION NUMBER: PCT/US01/08631
PRIOR FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: US 09/552,929
PRIOR FILING DATE: 2000-04-18
PRIOR APPLICATION NUMBER: US 09/770,160
PRIOR FILING DATE: 2001-01-26
PRIOR APPLICATION NUMBER: PCT/US01/08656
PRIOR FILING DATE: 2001-04-18
PRIOR APPLICATION NUMBER: US 09/577,408
PRIOR FILING DATE: 2000-05-18
PRIOR APPLICATION NUMBER: PCT/US01/14827
PRIOR FILING DATE: 2001-05-16
NUMBER OF SEQ ID NOS: 478
SOFTWARE: pf_genes Version 5.0
SEQ ID NO 175
LENGTH: 244
TYPE: PRT
ORGANISM: Homo sapiens
US-60-323-349-175

```

```

Query Match      85.5%; Score 1196; DB 24; Length 244;
Best Local Similarity 87.8%; Pred. No. 1.7e-116;
Matches 244; Conservative 0; Mismatches 0; Indels 34; Gaps 1;

QY 1 MHKAGLLGLCARAWNSVRMASSGTRDRPLANKVALVTASTDGIQFAIARRLAQDGAHV 60
Db 1 MHKAGLLGLCARAWNSVRMASSGTRDRPLANKVALVTASTDGIQFAIARRLAQDGAHV 60
QY 61 VSSRKQONVDQAVATLOGEGSLVTGTVCHVKGAEADRERLVATAVKLHGIDILVSNAAVN 120
Db 61 VSSRKQONVDQAVATLOGEGSLVTGTVCHVKGAEADRERLVAT----- 102
QY 121 PFGSIMDVTVEVMDKTLIDINVKAPALMTKAVPEMEKRRGGSVVIVSSIAAFSPGFS 180
Db 103 -----TLDINVKAPALMTKAVPEMEKRRGGSVVIVSSIAAFSPGFS 146
QY 181 PYNVSKTALLGLTKLTAIELAPRINRVNCLAPGLIKTSFSRMLWMDKEESMKETLIR 240
Db 147 PYNVSKTALLGLTKLTAIELAPRINRVNCLAPGLIKTSFSRMLWMDKEESMKETLIR 206

```

```

:
:
: TYPE: PRT
:
: ORGANISM: Homo sapiens
:
: FEATURE:
:
: NAME/KEY: SITE
:
: LOCATION: (6)
:
: OTHER INFORMATION: Xaa
:
: NAME/KEY: SITE
:
: LOCATION: (103)
:

```

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (127)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
PCT-US01-01239-1013

Query Match 64.8%; Score 906; DB 1; Length 193;
Best Local Similarity 68.0%; Pred. No. 3.4e-86;
Matches 189; Conservative 0; Mismatches 3; Indels 86; Gaps 1;
QY 1 MHRKAGLLGLCARAWNSVRMASSGTRDPLANKVALVTASTDGGIGFAIARRLAODGAHV 60
DB 2 MHRKAGLLGLCARAWNSVRMASSGTRDPLANKVALVTASTDGGIGFAIARRLAODGAHV 61
QY 61 VSSRKQONVDQAVATLOGEGSLVTGTVCHVKAEDRERLAVATVAKLHGGIDILVNAVN 120
DB 62 VSSRKQONVDQAVATLOGEGSLVTGTVCHVKAEDRERLAVATVAKLHGGIDILVNAVN 121
QY 121 PFGSMDVTEEVWVKTLDINVKAPALMTKAVVPMEKRGSGVIVSSIAAFSPGFS 180
DB 122 PFGSMDVTEEVWVKTLDINVKAPALMTKAVVPMEKRGSGVIVSSIAAFSPGFS 180
QY 181 PYNVSKTALLGLTKTLAIELAPRNVNCLAPGLIKTSFSRMLWMDKEESMKETLRIR 240
DB 182 PYNVSKTALLGLTKTLAIELAPRNVNCLAPGLIKTSFSRMLWMDKEESMKETLRIR 240
QY 241 RLGEPCAGIVSFLCSEDASYITGETVTVVGGGTPSRL 278
DB 242 RLGEPCAGIVSFLCSEDASYITGETVTVVGGGTPSRL 278
QY 156 RLGEPCAGIVSFLCSEDASYITGETVTVVGGGTPSRL 193
DB 157 RLGEPCAGIVSFLCSEDASYITGETVTVVGGGTPSRL 193

RESULT 14
PCT-US01-01308-456
Sequence 456, Application PC/TUS0101308
GENERAL INFORMATION:
APPLICANT: Human Genome Sciences, Inc., et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PJ204PCT
CURRENT APPLICATION NUMBER: PCT/US01/01308
CURRENT FILING DATE: 2001-01-14
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 495
SOFTWARE: PatentIn Ver 2.0
SEQ ID NO 456
LENGTH: 193
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (6)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (103)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (127)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
PCT-US01-01308-456

Query Match 64.8%; Score 906; DB 1; Length 193;
Best Local Similarity 68.0%; Pred. No. 3.4e-86;
Matches 189; Conservative 0; Mismatches 3; Indels 86; Gaps 1;
QY 1 MHRKAGLLGLCARAWNSVRMASSGTRDPLANKVALVTASTDGGIGFAIARRLAODGAHV 60
DB 2 MHRKAGLLGLCARAWNSVRMASSGTRDPLANKVALVTASTDGGIGFAIARRLAODGAHV 61
QY 61 VSSRKQONVDQAVATLOGEGSLVTGTVCHVKAEDRERLAVATVAKLHGGIDILVNAVN 120
DB 62 VSSRKQONVDQAVATLOGEGSLVTGTVCHVKAEDRERLAVATVAKLHGGIDILVNAVN 121

QY 121 PFGSMDVTEEVWVKTLDINVKAPALMTKAVVPMEKRGSGVIVSSIAAFSPGFS 180
DB 122 PFGSMDVTEEVWVKTLDINVKAPALMTKAVVPMEKRGSGVIVSSIAAFSPGFS 180
QY 181 PYNVSKTALLGLTKTLAIELAPRNVNCLAPGLIKTSFSRMLWMDKEESMKETLRIR 240
DB 182 PYNVSKTALLGLTKTLAIELAPRNVNCLAPGLIKTSFSRMLWMDKEESMKETLRIR 240
QY 241 RLGEPCAGIVSFLCSEDASYITGETVTVVGGGTPSRL 278
DB 242 RLGEPCAGIVSFLCSEDASYITGETVTVVGGGTPSRL 278
QY 156 RLGEPCAGIVSFLCSEDASYITGETVTVVGGGTPSRL 193
DB 157 RLGEPCAGIVSFLCSEDASYITGETVTVVGGGTPSRL 193

RESULT 15
PCT-US01-01349-673
Sequence 673, Application PC/TUS0101349
GENERAL INFORMATION:
APPLICANT: Human Genome Sciences, Inc., et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PJ206PCT
CURRENT APPLICATION NUMBER: PCT/US01/01349
CURRENT FILING DATE: 2001-01-14
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 939
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 673
LENGTH: 193
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (6)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (103)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (127)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
PCT-US01-01349-673

Query Match 64.8%; Score 906; DB 1; Length 193;
Best Local Similarity 68.0%; Pred. No. 3.4e-86;
Matches 189; Conservative 0; Mismatches 3; Indels 86; Gaps 1;
QY 1 MHRKAGLLGLCARAWNSVRMASSGTRDPLANKVALVTASTDGGIGFAIARRLAODGAHV 60
DB 2 MHRKAGLLGLCARAWNSVRMASSGTRDPLANKVALVTASTDGGIGFAIARRLAODGAHV 61
QY 61 VSSRKQONVDQAVATLOGEGSLVTGTVCHVKAEDRERLAVATVAKLHGGIDILVNAVN 120
DB 62 VSSRKQONVDQAVATLOGEGSLVTGTVCHVKAEDRERLAVATVAKLHGGIDILVNAVN 121
QY 121 PFGSMDVTEEVWVKTLDINVKAPALMTKAVVPMEKRGSGVIVSSIAAFSPGFS 180
DB 122 PFGSMDVTEEVWVKTLDINVKAPALMTKAVVPMEKRGSGVIVSSIAAFSPGFS 180
QY 181 PYNVSKTALLGLTKTLAIELAPRNVNCLAPGLIKTSFSRMLWMDKEESMKETLRIR 240
DB 182 PYNVSKTALLGLTKTLAIELAPRNVNCLAPGLIKTSFSRMLWMDKEESMKETLRIR 240
QY 241 RLGEPCAGIVSFLCSEDASYITGETVTVVGGGTPSRL 278
DB 242 RLGEPCAGIVSFLCSEDASYITGETVTVVGGGTPSRL 278
QY 156 RLGEPCAGIVSFLCSEDASYITGETVTVVGGGTPSRL 193
DB 157 RLGEPCAGIVSFLCSEDASYITGETVTVVGGGTPSRL 193

Search completed: January 22, 2002, 15:21:30
Job time: 185 sec

1	618.5	44.2	253	5	US-09-708-427-59234	Sequence 59234, A
2	612.5	43.8	253	5	US-09-708-427-52797	Sequence 52797, A
3	575.5	41.1	234	5	US-09-708-427-53798	Sequence 52798, A
4	377.5	27.0	263	5	US-09-815-242-14092	Sequence 14092, A
5	368.5	26.3	283	5	US-09-815-242-10241	Sequence 10241, A
6	353	25.2	264	5	US-09-708-427-17916	Sequence 17916, A
7	352	25.2	273	5	US-09-708-427-51897	Sequence 51897, A
8	348	24.9	246	5	US-09-815-242-5461	Sequence 5461, A
9	348	24.9	246	5	US-09-815-242-12123	Sequence 12123, A
10	348	24.9	246	5	US-09-815-242-12803	Sequence 12803, A
11	348	24.9	246	5	US-09-815-242-13100	Sequence 13100, A
12	348	24.7	251	5	US-09-572-810A-2	Sequence 2, Appl
13	346	24.7	251	5	US-09-815-242-5689	Sequence 5689, Ap
14	346	24.7	272	5	US-09-815-242-12688	Sequence 12688, A
15	345.5	24.7	269	5	US-09-708-427-17881	Sequence 17881, A
16	345	24.7	253	5	US-09-708-427-17917	Sequence 17917, A
17	344.5	24.6	203	5	US-09-634-955B-1	Sequence 20, Appl
18	341	24.4	268	5	US-09-708-427-17920	Sequence 20, Appl
19	340.5	24.3	147	5	US-09-708-427-59235	Sequence 59235, A
20	340	24.3	242	5	US-09-815-242-10974	Sequence 10974, A
21	340	24.3	243	5	US-09-815-242-13360	Sequence 13360, A
22	340	24.3	243	5	US-09-815-242-13581	Sequence 13581, A
23	338.5	24.2	147	5	US-09-708-427-52799	Sequence 52799, A
24	338.5	24.2	203	5	US-09-634-955B-26	Sequence 26, Appl
25	336.5	24.1	203	5	US-09-634-955B-17	Sequence 17, Appl
26	332.5	23.8	250	5	US-09-708-427-17882	Sequence 17882, A

RESULT 2

US-09-708-427-52797

; Sequence 52797, Application us/09708427

; GENERAL INFORMATION:

; APPLICANT: N. ALEXANDROV et al.

; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES

; FILE REFERENCE: 2750-1243P

; CURRENT APPLICATION NUMBER: US/09/708,427

; CURRENT FILING DATE: 2000-11-09

; NUMBER OF SEQ ID NOS: 85364

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 52797

; LENGTH: 253

; TYPE: PRT

; ORGANISM: zea mays subsp. mays

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: 1..253

; OTHER INFORMATION: Xaa is any amino acid

; NAME/KEY: misc_feature

; LOCATION: 1..253

; OTHER INFORMATION: Ceres Seq. ID 1929770

US-09-708-427-52797

```

Query Match      43.8%  Score 612.5  DB 5  Length 253;
Best Local Similarity 51.8%  pred. No. 2.4e-48;
Matches 129;  Conservative 39;  Mismatches 78;  Indels 3;  Gaps 2;

QY 30 LANKVALVTASTDIGFAIAHRLAQQDGAHVVSRRKQONVDOAVATLQGEGLSVGTGYCH 89
    I  |||:||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8  LEQKVAIVTASTMIGLIGLAEERLGLEGAHVVISRRKQKNVXAVEGLRAKGITAVGAVCH 67

QY 90 VGRAEDRERLVATAKLHGGDILVSNVAVNPPFGSMDVITEEVDKDTLDINVKAPALMT 149
    I  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 68 VSDAQQRKSIETAVKSGFGHIDILVSNAAANPVSVDILEMESVLDKLDKWDINVKASILI 127

QY 150 KAVVPEMEKRGGSWTVSSIAAFSPSPGFSPYNNVKTALLGNTKTLAEIAPRNIIVNC 209
    -  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 128 QDAAPHLRK--GSSVLIISSTAGVNPQGLTMYXVTKTALFGLTKALAGEMP-DTRVNC 184

QY 210 LAPGLIKTFSRMLMDKKEESMKETIRIRRLGEPEDCAGIVSPGLSEDASYITGETW 269
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 185 IXPGEVPTFRASFLTENETIRKELNERTKLRGLCTVEDMAAAAFASDDASYITAETIV 244

QY 270 VGGGTFSRL 278*
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 245 VAGGVOSRL 253
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 3
 US-09-708-427-52798
 ; Sequence 52798, Application US/09708427
 ; GENERAL INFORMATION:
 ; APPLICANT: N. ALEXANDROV et al.
 ; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTIDES
 ; TITLE OF INVENTION: THEREBY
 ; FILE REFERENCE: 2750-1243P
 ; CURRENT APPLICATION NUMBER: US/09/708,427
 ; CURRENT FILING DATE: 2000-11-09
 ; NUMBER OF SEQ ID NOS: 85364
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 52798
 ; LENGTH: 234
 ; TYPE: PRT
 ; ORGANISM: Zea mays subsp. mays
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: 1..234
 ; OTHER INFORMATION: Xaa is any amino acid

```
; NAME/KEY: misc_feature
; LOCATION: 1..234
; OTHER INFORMATION: Ceres Seq. ID 1929771
US-09-708-427-52798

Query Match      41.1%; Score 575.5; DB 5; Length 234;
Best Local Similarity 50.8%; Pred. No. 5.1e-45;
Matches 120; Conservative 38; Mismatches 75; Indels 3; Gaps

QY   43  GIGFATARLLAQDGAHVVSSRKQNVDOAVATLQEGLSVTCTVCHVGKAEDRRLVAT 102
    ||| ||| ||| ||| ||| ||| ||| ||| : ||| : ||| : ||| : ||| : |||
Db   .2  GIGLATAERLGLEAGAAVVISRRKKNNXAVEGIRAKGITAVGAVCHVDQAQRKSLIET 61
    ||| ||| ||| ||| ||| ||| ||| ||| : ||| : ||| : ||| : ||| : |||

QY   103 AVKLHGHDITLVNNAANPPFGSIMDVTEVBWKTIDINVKAPALMTKAVVPMEKRGG 162
    ||| ||| ||| ||| ||| ||| ||| ||| : ||| : ||| : ||| : ||| : |||
Db   62  AVKSFGHIDILVNRAANPSVDILEMKESVLDKLNWINVKASILIQDAAPHLRK--GS 119

QY   163 SVVIVSIAAFSPGSPFPYNVSKTALLGLTKTLAIELAPRNTRVNCPLAPGLIKTFSRM 222
    ||| : ||| ||| : ||| ||| ||| ||| ||| : ||| : ||| : ||| : ||| : |||
Db   120 SVIIISIIAGNPQQGLTMYYVTKTALFGLTKALAGEMGP-DTRVNCIXPGFPVPTRFASF 178

QY   223 LWNMDKEEESMKETLIIRRUGPEDCAGIVSFCCSEDASYITGETVTVGGGTSPRL 278
    ||| : ||| : ||| ||| ||| ||| ||| ||| : ||| : ||| : ||| : ||| : |||
Db   179 LTNETIRKELNERTLKRGUVEDMAAAAFAASDDASYTAETITWAGGVQSRL 234
```

```

RESULT      4
US-09-815-242-14092
; Sequence 14092, Application US/09815242
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zsaskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA 011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14092
; LENGTH: 263
; TYPE: PNT
; ORGANISM: Salmonella typhi
US-09-815-242-14092

```

```
Query Match          27.0%; Score 377.5; DB 5; Length 263;
Best Local Similarity 35.3%; Pred. No. 7.le-27;
Matches 89; Conservative 53; Mismatches 101; Indels 9; Gaps 4;
```

Db 4 LTGKTALITGASQIGEGIAVFARFAGANLILLDISDE-IEKLADELGGRHRCRTAVKAD 62
Qy 90 VGKAEDRERLVATVYKLGHGIDILVSNAAVNPFFGSGIMDVTDKLDINVKAPALMT 149
Db 63 VRDFASVQAAVARAKETGRDILVNNAGVCR-LGNFLDMSSEDRDFHIDINIKGVNVT 121
Qy 150 KAVYPENKRGSGSVVIVSSIAA-FSPSPGSPYNVSKTALLGTTKTLAIELAPRINRVN 208
Db 122 KAVLPEMIKRRDGRIVNMSSVTGDMVADRGETAYALSKAAIVGLTKSLAVEYAQSGIRVN 181
Qy 209 CLAPGLIKTSFRLMWMDEKEE-----SMKETLRIRRLGPEPCAGIVSFLCSEDASY 262
Db 182 AICPGYVTPMAESIAQSNDDPDESULTEMAKAIPRLADPLEVGEAFLASDESSY 241
Qy 263 ITGETVVVGGGT 274
Db 242 LTGTQNVIDGS 253

RESULT 5
US-09-815-242-10241
; Sequence 10241, Application US/09815242
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; TITLE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA-011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10241
; LENGTH: 285
; TYPE: PRT
; ORGANISM: Escherichia coli
US-09-815-242-10241

Query Match 26.3%; Score 368.5; DB 5; Length 285;
Best Local Similarity 32.6%; Pred. No. 5.3e-26;
Matches 88; Conservative 56; Mismatches 117; Indels 9; Gaps 4;
Qy 12 RANVSVMASGMRTRDLANKVALVASTDGGIGFAIRLAQDGAHVYSSRRKQNVQDQ 71
Db 8 REWSVTLCINKESIKMKGLKGTALITGALQIGEGITARTFARHANLILLDISPE-IEK 66
Qy 72 AVATLQEGISVTGTVCHVGAEDRERLVATVYKLGHGIDILVSNAAVNPFFGSGIMDVTE 131
Db 67 LADELGGRHRCRTAVADVDRDPASVAAIKRAKEGREGIDILVNNAGVCR-LGSFLDMSD 125
Qy 132 EVWDKTLDINVKAPALMTKAVVPEMEKRGSGSVVIVSSIAA-FSPSPGSPYNVSKTALL 190

Db 126 DDRDFHIDINIKGVNVTKAVLPEMIARKDGRIVNMSSVTGDMVADPGETAVALTKAAIV 185
Qy 191 GLTKTALATEAPRINRVNCLAPGLIKTSFRLMWMDEKEE-----SMKETLRIRRLGE 244
Db 186 GLTSLAVEYAQSGIRVNAICPGYVTPMAESIAQSNDDPDESULTEMAKAIPMRRLAD 245
Qy 245 PEDCAGIVSFLCSEDASYITGETVVVGGGT 274
Db 246 PLEVGEAFLASDESSYLTGTQNVIDGS 275

RESULT 6
US-09-708-427-17916
; Sequence 17916, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; TITLE OF INVENTION: THEREBY
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 17916
; LENGTH: 262
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..262
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..262
; OTHER INFORMATION: Ceres Seq. ID 1833207
US-09-708-427-17916

Query Match 25.2%; Score 353; DB 5; Length 262;
Best Local Similarity 35.0%; Pred. No. 1.2e-24;
Matches 89; Conservative 47; Mismatches 110; Indels 8; Gaps 3;
Qy 24 MTRDPLANKVALVYASTDGGIGFAIRLAQDGAHVYSSRRKQNVQDGAVALQEGLSV 83
Db 1 MDKRWSLKGMTALVTGGASGIGYAIVEELAGFGARIHVCIDISEAKLNSLSEWEKKGQV 60
Qy 84 TGTVCHVGAEDRERLVAT-AVKLHGIDILVSNAAV---NPFFGSGIMDVTEEVWDKTLD 139
Db 61 SGSCVDVASRPEREELMOTVSSQFDGKLNILVSNVGVIRSKP---TTEYTEDDDFAFHIS 116
Qy 140 INVKAPALMTKAVVPEMEKRGSGSVVIVSSIAAFSPSPGSPYNVSKTALLGLTKTLAIE 199
Db 117 SNVEAAVHFSQLSHPLLKASGYSIIFVSSIAGVISFDAGSIYGLTKGALIQALKNLACE 176
Qy 200 LAPRINRVNCLAPGLIKTSFRLMWMDEKESMKETLRIRRLGPEPCAGIVSFLCSED 259
Db 177 WAKGCIANAVAPNVINTPLSQSYLEDVSFKALLSRTPGLRGVGEPEVASLVAFCLPLA 236
Qy 260 ASYITGETVVVGGG 273
Db 237 ASYITGOTICVDGG 250

RESULT 7
US-09-708-427-51897
; Sequence 51897, Application US/09708427
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: SEQUENCE-DETERMINED DNA FRAGMENTS AND CORRESPONDING POLYPEPTID
; TITLE OF INVENTION: THEREBY
; FILE REFERENCE: 2750-1243P
; CURRENT APPLICATION NUMBER: US/09/708,427
; CURRENT FILING DATE: 2000-11-09
; NUMBER OF SEQ ID NOS: 85364

SOFTWARE: PatentIn version 3.1
SEQ ID NO 51897
LENGTH: 273
TYPE: PRT
ORGANISM: Zea mays subsp. mays
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..273
OTHER INFORMATION: Xaa is any amino acid
NAME/KEY: misc_feature
LOCATION: 1..273
OTHER INFORMATION: Ceres Seq. ID 1927824
US-09-708-427-51897

Query Match 25.2%; Score 352; DB 5; Length 273;
Best Local Similarity 34.1%; Pred. No. 1.6e+24;
Matches 86; Conservative 49; Mismatches 115; Indels 2; Gaps 2;
Qy 23 GWTREDPLANKVALVTASTDGIGFAIARRLAQDGAHVVSRRKQONVDQAVATLQGEGLS 82
Db 12 GSSGKWLHGTALVTGTRGIGRAVVEELAAALGAHVTCRKAEEELGERIKWEARGFS 71
Qy 83 VTGTVCHVGKAEEDRLVATAVKLHGG-IDILVSNAAVNPFFGSDIMVTEEVWDTLDIN 141
Db 72 VTGVCDSLSEDRQLREVRADRFGGKLNILVNVGTN-IRKPTTETAEYFSLMATN 130
Qy 142 VKAPALMTKAVVPEMEKRGGSVVIVSSIAAFSPGSPYNVSKTALLGLTKTLAIELA 201
Db 131 LESAYHLCQIAHPLLLKSGSGSIIFISSVAGAIGFSCTIYAMTKGAINQLTKNLACWA 190
Qy 202 PNIRVNCVLAFLIKTSFRLMWDKKEESKMETLRRRLGEPEDCAGIVSFLCSESDAS 261
Db 191 KDNIRANSVAFWYITLSTEGILANKNFEEQVSRTPILGRVGRGEPGEVSALVFLCMPGST 250
Qy 262 YITGETVVVGGG 273
Db 251 YISQGTIAVDGG 262

RESULT 8
US-09-815-242-5461
Sequence 5461, Application US/09815242
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
PRIOR FILING DATE: 2001-03-21
PRIOR FILING DATE: 2000-03-21
PRIOR FILING DATE: 2000-05-23
PRIOR FILING DATE: 2000-05-26
PRIOR FILING DATE: 2000-10-23
PRIOR FILING DATE: 2000-11-27
PRIOR FILING DATE: 2000-12-22
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSEQ for Windows Version 4.0

SEQ ID NO 5461
LENGTH: 246
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-5461

Query Match 24.9%; Score 348; DB 5; Length 246;
Best Local Similarity 32.6%; Pred. No. 3.2e+24;
Matches 79; Conservative 57; Mismatches 102; Indels 4; Gaps 3;
Qy 33 KVALVTASTDGIGFAIARRLAQDGAHVVS-SRKQONVDQAVATLQGEGLSVTGTVCVHG 91
Db 5 KVALVTGASRGIGRSIALQLAEBEGYNVAVNYAGSKAEAVVEEIKAGVDSFAIQANVA 64
Qy 92 KAEDRERLVAATAVKLHGGIDILVSNAAVNPFFGSDIMVTEEVWDTLDINVKAPALMTKA 151
Db 65 DADEVKAMIKVVSQFSLDLVNVNAGITR-DNLLMRKKEQEDVDVDTNKLGVFNCIOK 123
Qy 152 VVPEMEKRGGSVVIVSSIAAFSPGSPYNVSKTALLGLTKTLAIELAPRNRVNCVLA 211
Db 124 ATPQMLRSGAIINLSSVGVAGVNGPOANVAVTKAGVIGLTKSAARELASRGITVNAVA 183
Qy 212 PGLIKTSFRLMWDKKEESKMETLRRRLGEPEDCAGIVSFLCSESDASVITGETVVVG 271
Db 184 PGFVSDMTDAL--SDELKEQMLTQIPLAREGQDPTDIANTVAFLASDRKAKYITGOTIHN 241
Qy 272 GG 273
Db 242 GG 243

RESULT 9
US-09-815-242-12123
Sequence 12123, Application US/09815242
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
PRIOR FILING DATE: 2001-03-21
PRIOR FILING DATE: 2000-03-21
PRIOR FILING DATE: 2000-05-23
PRIOR FILING DATE: 2000-05-26
PRIOR FILING DATE: 2000-10-23
PRIOR FILING DATE: 2000-11-27
PRIOR FILING DATE: 2000-12-22
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSEQ for Windows Version 4.0
SEQ ID NO 12123
LENGTH: 246
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-12123

Query Match 24.9%; Score 348; DB 5; Length 246;


```

Db      184 PGFIVSDMTDAL--SDELKEQMTCIPLARFGDQDTLANTVAELASDKAKYITGQTHVN 241
QY      272 GG 273
Db      242 GG 243

```

RESULT 12

```

US-09-572-810A-2
; Sequence 2, Application US/09572810A
; GENERAL INFORMATION:
; APPLICANT: Huang, Jianzhong
; APPLICANT: McDevitt, Damien
; TITLE OF INVENTION: Fabg A
; FILE REFERENCE: GM10192
; CURRENT APPLICATION NUMBER: US/09/572,810A
; CURRENT FILING DATE: 2000-05-17
; PRIOR APPLICATION NUMBER: 09/238,481
; PRIOR FILING DATE: 1999-01-28
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0

```

; CLN ID NO 2
 ;
 ; LENGTH: 246
 ;
 ; TYPE: PRT
 ; ORGANISM: Staphylococcus aureus
 US-09-572-810A-2

Query Match 24.9%; Score 348; DB 5; Length 246;
Best Local Similarity 32.6%; Pred. No. 3.2e-24;
Matches 79; Conservative 57; Mismatches 102; Indels 4; Gaps 3;

Qy - 33 KVALVTASTDGIGFAIARLQAQGAHVWS-SRKQQNVDAQAVATLOGEGLSVTGTCHVG 91
| | | | : : | | : : | | : : | | : : | | : : | | : : |
Db 5 K\$ALVTCASRGIRSIQLAEEGVNAVNTAGSKEKAEVVTEIKAKGVDSFATQANVA 64

Oy	92	K A E D R E R L V A T A V K L H G G I D I L S N A A V N P F G S I M D V T E E V W D K T L D I N V K A P A L M T K A	151
		: : : : : : : : : : : : : : : : : : : : : : :	
Dd	65	D A D E V K A M I K E V V S Q F G S L D V L V N N A G I T R - D N L L M R K E Q E W D D V I D T N L K G V E N C I O Q S	123

Qy 152 VPPEMKRGGGVIVSSIAAFSPGFPYVNVSKTALLGLTKLTALIELAPRNIRVNCLA 211
 :|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::
 Db 124 ATPQMLRQRSGAIIINLSSVVGAVNGPQANVTAKCVIGLYKSAAREIASRGITVNAVA 183

QY 212 PGLIKTSFRLMWDKEXESMKETLIRRIERGEPEDCAGIVSFLCSEDASYITGETVVVG 271
z ||| : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : |
Db 184 PGFVSDMTDAL--SDLEKEQLTQIPLARFGQDTDIAVTAVFLASDRAKYITGOTIHVN 241

Qy	272	GG	273
		==	
Db	242	GG	243

RESULT 13

US-09-815-242-5689
 : Sequence 5689, Application US/09815242
 : GENERAL INFORMATION:
 : APPLICANT: Hasebeck, Robert
 : APPLICANT: Ohlsen, Kari L.
 : APPLICANT: zyskind, Judith W.
 : APPLICANT: Wall, Daniel
 : APPLICANT: Trawick, John D.
 : APPLICANT: Carr, Grant J.
 : APPLICANT: Yamamoto, Robert T.
 : APPLICANT: Xu, H. Howard
 : TITLE OF INVENTION: Identification of
 : TITLE OF INVENTION: Prokaryotes
 : FILE REFERENCE: ELITRA.011A
 : CURRENT APPLICATION NUMBER: US/09/815
 : CURRENT FILING DATE: 2001-03-21
 : PRIOR APPLICATION NUMBER: 60/191,078
 : PRIOR FILING DATE: 2000-03-21

```

; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ. ID NOS: 14110
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 5689
; LENGTH: 251
; TYPE: prt
; ORGANISM: Staphylococcus aureus
US-09-815-242-5689

```

Query Match	24.7%;	Score 346;	DB 5;	Length 251;
Best Local Similarity	33.6%;	Pred. No. 5e-24;		
Matches 84;	Conservative 56;	Mismatches 102;	Indels 8;	Gaps 4;

Qy	30	LANKVALVTASTDGGFATARRLQAQAHVVSSRRQQNVDAVATLQCEGLSVTCVTCH	89
Db	1	LENKVAIVTGASTGICGASAIALAQGAQVLAVDIAEA-VSETVDIKKNGDKAKAYVD	59
Qy	90	VGRAEDRELRVATAVKLHGIDILTVNAAVNPFFGIMDVTVEWPKTILIDINVKAPALMT	149
Db	60	IASEQIDNFASIERQFGHVDVLFNAGVDNAAGRIHEVPTDVKIINVDMRGFTLMT	119

[illegible]

QY 210 LAPGLIKTSF-SRMLWMDKEKE-----ESMKETLRIRRLGPEPCDAGIVSFLCSDEDASYI 263
 :|||:||| :::: :|||:||| :|||::::|
Db 179 IAPGTITPLVDKLTGTSEDKEGAIFRENOCKWMPTLGRLGKPBEVGKLXVFLASDESSFI 238

```

Qy      264 TGETVVVGGG 273
        ||||: ||
Db      239 TGETIRIDGG 248

```

RESULT 14

US-09-815-242-13688
Sequence 12688, Application US/09815242
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
TITLE OF INVENTION: Prokaryotes
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27

QY 75 TLQEGLSVTGTVCHVGKAEDRERLVATAVKL-HGGIDILVSNAAVNPFEGSIMDVTE--131

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2002, 15:15:50 ; Search time 23.33 seconds
(without alignments)
882.656 Million cell updates/sec

Title: US-09-866-034-2

Perfect score: 1399

Sequence: 1 MHKAGLLGLCARAWNSVRMA.....DASYITGETVVVGGTSPSL 278

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

A_Geneseq_1101.*
1: /SID88/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SID88/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SID88/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SID88/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SID88/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SID88/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SID88/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SID88/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SID88/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SID88/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SID88/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SID88/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SID88/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SID88/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SID88/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SID88/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SID88/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SID88/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SID88/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SID88/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SID88/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SID88/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1399	100.0	278	21	PRO1800, a Hep27 h
2	1395	99.7	278	22	Human protein sequ
3	1327	94.9	278	21	Short chain alcoho
4	649.5	46.4	183	21	Breast and ovarian
5	588.5	42.1	254	21	Arabidopsis thalia
6	588.5	42.1	276	21	Arabidopsis thalia
7	428	30.6	103	22	Human colon cancer
8	404	28.9	248	21	Amino acid sequenc
9	370	26.4	79	21	Human secreted pro
10	353	25.2	262	21	Arabidopsis thalia
11	349	24.9	292	22	C glutamicum prote

12	349	24.9	295	22	AA079400	Corynebacterium g)
13	349	24.9	295	22	AA079401	Corynebacterium g)
14	348	24.9	246	21	AA015707	Staphylococcus aur
15	348	24.9	246	22	AA020195	S. aureus NADPH-de
16	345	24.7	253	21	AA031819	Arabidopsis thalia
17	341	24.4	243	22	AA010332	CFE 35 protein seq
18	340	24.3	243	19	AA080670	S. pneumoniae fatt
19	340	24.3	243	21	AA015706	Streptococcus pneu
20	339	24.2	254	22	AA049773	Protein with acety
21	339	24.2	306	22	AA081644	S. epidermidis ope
22	336	24.0	248	20	AA036959	Chlamydia trachoma
23	333.5	23.8	251	20	AA034891	Chlamydia pneumoni
24	333.5	23.8	262	22	AA083032	S. epidermidis ope
25	331.5	23.7	261	21	AA096271	B. subtilis glucos
26	331.5	23.7	261	21	AA054424	Amino acid sequenc
27	331.5	23.7	263	21	AA032504	Arabidopsis thalia
28	330	23.6	273	21	AA043271	Arabidopsis thalia
29	330	23.6	273	21	AA043270	Arabidopsis thalia
30	328.5	23.5	315	17	AA089323	Rape leaf beta-ket
31	328.5	23.5	315	17	AA089322	Rape seed beta-ket
32	327.5	23.4	272	21	AA010740	B. megaterium gluc
33	327.5	23.4	340	21	AA010741	H. ghilianii/B. me
34	326.5	23.3	261	13	AA027757	Glucose dehydrogen
35	326	23.3	279	21	AA049315	Arabidopsis thalia
36	325.5	23.3	261	9	AA080590	Sequence of glucos
37	325.5	23.3	261	13	AA027756	NAD affinity gluc
38	324	23.2	282	21	AA044578	Xylitol dehydrogen
39	324	23.2	262	22	AA086337	G. suboxydans DSM
40	324	23.2	271	21	AA080828	Arabidopsis thalia
41	323.5	23.1	254	21	AA032505	Arabidopsis thalia
42	323.5	23.1	264	21	AA049317	Arabidopsis thalia
43	323.5	23.1	269	21	AA049316	Arabidopsis thalia
44	323.5	23.1	318	21	AA029313	Arabidopsis thalia
45	323.5	23.1	322	21	AA029312	Arabidopsis thalia

ALIGNMENTS

RESULT 1

AA096729

ID AA096729 standard; Protein; 278 AA.

XX

AC AA096729;

XX 26-SEP-2000 (first entry)

DT PRO1800, a Hep27 homologue.

DE

XX PRO1800; Hep27; homologue; short-chain alcohol dehydrogenase; SCAD;

KW secreted protein; transmembrane protein; recombinant production;

KW gene therapy.

XX

OS Homo sapiens.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX WO200036102-A2.
XX 22-JUN-2000.
XX 01-DEC-1999; 99WO-US28634.
XX 16-DEC-1998; 98US-0112851.
XX 16-DEC-1998; 98US-0113145.
XX 22-DEC-1998; 98US-0113511.
XX 12-JAN-1999; 99US-0115558.
XX 12-JAN-1999; 99US-0115565.
XX 12-JAN-1999; 99US-0115733.
XX 09-FEB-1999; 99US-0119341.
XX 10-FEB-1999; 99US-0119341.
XX 12-FEB-1999; 99US-0119341.
XX 02-JUN-1999; 99WO-US12252.
XX (GETH) GENENTECH INC.
XX Botstein D, Desnoyers L, Ferrara N, Fong S, Gao W, Goddard A;
PI Gurney AL, Pan J, Roy MA, Stewart TA, Tamas D, Watanabe CK;
PI Wood W;
XX WPI; 2000-431586/37.
XX N-PSDB; AAA51259.
XX Isolated nucleic acid molecule encodes a PRO polypeptide which is a
XX transmembrane polypeptide
XX Claim 12; Fig 2; 154pp; English.
XX This is PRO1800, a putative human Hep27 protein homologue, which has
XX recently been shown to have homology to known short-chain alcohol
XX dehydrogenase (SCAD) family of proteins. The invention concerns novel
XX secreted and transmembrane proteins, designated PRO polypeptides. The
XX cDNA and gene sequences are useful in the recombinant production of PRO
XX polypeptides, as a hybridization probe to screen libraries to isolate
XX cDNAs with sequence identity to PRO polypeptides or to map the gene
XX encoding the PRO polypeptides and analyzing genetic disorders. The
XX cDNA/gene can also be used to produce transgenic animals useful for the
XX development and screening of therapeutically useful reagents. They can
XX also be used in gene therapy, e.g. to replace a defective gene.
XX Sequence 278 AA;
Query Match 100.0%; Score 1399; DB 21; Length 278;
Best Local Similarity 100.0%; Pred. No. 4.9e-132;
Matches 278; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MHKAGLLGLCARAWNSVRVSSGMRDRPLANKVALVTASTDGTGFAIARRLAQDGAHV 60
Db 1 mhkagllglcarawnsvrmasgmtrdrplankvalvtastdgtgfaiarlraqdghv 60
QY 61 VSSRQNDVQAVATLQGLSVGTGCHVKAEDRERLVAFAVKLHGIDILVNAVN 120
Db 61 vsrqnqndvqavatlqgelsvgtgchvkaedrervlvaavklhgidilvnaavn 120
QY 121 PFFGSTMVTEEWKDTLDINVKAPALMTKAVVPEMEKRGGSVVIVSSTAATSPSPGFS 180
Db 121 pffgsimdvteewkdtldinvkapalmckavvpemekrggsvvsvsstaatspspgfs 180
QY 181 PYNVSKTALLGLTKTLATLAPRNIRVNCAPGLIKTSFMRMLWMDKEESMKETLIR 240
Db 181 pynvsktallgltktlatlaprnirvncapglktsfmrmlwmdkeesmketlir 240
QY 241 RLGEPEDCAGIVFLCSDEASVITGETVVVGGTSPRL 278
Db 241 rlgepedcagivflcsedasyitgetvvvvggtpsl 278

RESULT 2

AAB93414
ID AAB93414 standard; Protein; 278 AA.
XX
AC AAB93414;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human protein sequence SEQ ID NO:12620.
XX
KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-0116126.
XX
PR 29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
DR WPI; 2001-318749/34.
XX
PS Claim 8; SEQ ID 12620; 2537pp + CD ROM; English.
XX
CC The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
SQ Sequence 278 AA;
Query Match 99.7%; Score 1395; DB 22; Length 278;
Best Local Similarity 99.6%; Pred. No. 1.2e-131;
Matches 277; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MHKAGLLGLCARAWNSVRVSSGMRDRPLANKVALVTASTDGTGFAIARRLAQDGAHV 60
Db 1 mhkagllglcarawnsvrmasgmtrdrplankvalvtastdgtgfaiarlraqdghv 60

QY 61 VSSRKQNVDOAVATLOGEGLSVTGTVCHVCKAEDRELRLVATVAVKLHGIDILVSNAAVN 120
Db 61 vssrkqnvddqavatlqgeglsvtgtvchvgkaedrelrlvatavklhgidilvsnaavn 120
QY 121 PFFGSIIMDVTEEVWDKTLIDINVKAPALMTKAVVPEMEKRGGSVVIVSSIAAFSPSPGFS 180
Db 121 pffgsimdvteevwkdltidinvkapalmckavvpemekrggsvvsvssiaafspspgfs 180
QY 181 PYNVSKTALLGLTKTLAELAPRINRVNCLAPGLIKTFSRMLWMDKEKESMKETLIR 240
Db 181 pynvsktallgltktlalelaprinvnclapglktsfsmwmdkekesmktetlir 240
QY 241 RLGPEDCAGIVSFLCSEDASYIGETVVVGGTSPRL 278
Db 241 rlgedcagivsfllcsedasyitgetvvvggtsprrl 278

RESULT 3
AAV68735
ID AAV68735 standard; Protein; 278 AA.
AC AAV68735;
XX
DT 05-MAY-2000 (first entry)
XX
DE Short chain alcohol dehydrogenase-related molecule ScRM-1 protein.
XX
KW Human; short chain alcohol dehydrogenase-related molecule;
KW SCAD-related molecule ScRM-1; SCRM-2; metabolic regulator;
KW cell proliferation regulator; inflammation regulator;
KW cell proliferative disorder; immune disorder; arteriosclerosis;
KW atherosclerosis; bursitis; cirrhosis; hepatitis; AIDS;
KW Addison's disease; adult respiratory distress syndrome; allergy;
KW ankylosing spondylitis; amyloidosis; cancer.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Modified-site 16
FT /note= "potential protein kinase C phosphorylation site"
FT Modified-site 21
FT /note= "potential protein kinase C phosphorylation site"
FT Modified-site 25
FT /note= "potential casein kinase II phosphorylation site
and potential protein kinase C phosphorylation
site"
FT Modified-site 31
FT /note= "potential protein kinase C phosphorylation site"
FT Region 33..45
FT /note= "SCAD motif"
FT Region 34..51
FT /note= "glucose/ribitol dehydrogenase motif"
FT Binding-site 39..46
FT /note= "AMP-binding domain"
FT Modified-site 62
FT /note= "potential protein kinase C phosphorylation site"
FT Modified-site 63
FT /note= "potential protein kinase C phosphorylation site"
FT Region 108..119
FT /note= "potential protein kinase C phosphorylation site"
FT Region 108..118
FT /note= "glucose/ribitol dehydrogenase motif"
FT Region 108..119
FT /note= "SCAD motif"
FT Modified-site 125
FT /note= "potential casein kinase II phosphorylation site"
FT Region 156..172
FT /note= "glucose/ribitol dehydrogenase motif"
FT Region 162..179
FT /note= "SCAD motif"
FT Region 162..173
FT /note= "SCAD signature sequence"
FT Region 169..197

FT Region /note= "SCAD family signature sequence"
FT 182..201
FT /note= "glucose/ribitol dehydrogenase motif"
FT Domain 182..186
FT /note= "canonical catalytic site of SCADS"
FT Region 203..220
FT /note= "glucose/ribitol dehydrogenase motif"
FT Region 204..213
FT /note= "SCAD motif"
FT Modified-site 232
FT /note= "potential casein kinase II phosphorylation site
and potential protein kinase C phosphorylation
site"
FT Modified-site 236
FT /note= "potential protein kinase C phosphorylation site"
FT Region 238..258
FT /note= "glucose/ribitol dehydrogenase motif"
XX WO200004135-A2.
XX
XX 27-JAN-2000.
XX
XX 16-JUL-1999; 99WO-US16164.
XX
XX 16-JUL-1998; 98US-0116750.
XX 16-JUL-1998; 98US-0160074.
XX
XX (INCY-) INCYTE PHARM INC.
XX
XX Bandman O, Tang YT, Corley NC, Azimzai Y, Baughn MR;
XX
XX WPI: 2000-171266/15.
XX N-PSDB: AAZ46090.
XX
XX New short chain alcohol dehydrogenase polypeptides useful for
XX diagnosis, treatment and prevention of cell proliferative disorders
XX such as atherosclerosis, cirrhosis and cancers of various tissues
XX
XX Claim 1; Fig 1A-D; 78pp; English.
XX
XX The present sequence represents a human short chain alcohol dehydrogenase
XX (SCAD)-related molecule designated SCRM-1. The specification also
XX describes SCRM-2. SCRM proteins are metabolic, cell proliferation and
XX inflammation regulators. The SCRM polynucleotides and polypeptides are
XX used for treating or preventing a cell proliferative or immune disorder
XX in humans. Cell proliferative disorders include arteriosclerosis,
XX atherosclerosis, bursitis, cirrhosis, and hepatitis. Immune disorders
XX include AIDS, Addison's disease, adult respiratory distress syndrome,
XX allergies, ankylosing spondylitis, and amyloidosis. The vectors,
XX agonists, antagonists, antibodies and complementary sequences are also
XX used for treating the above conditions. The polynucleotides and
XX polypeptides are also used for treating cancers of various tissues
XX such as adrenal gland, bladder, bone, bone marrow, and brain.
XX
XX Sequence 278 AA;
XX

Query Match 94.9%; Score 1327; DB 21; Length 278;
Best Local Similarity 96.4%; Pred. No. 8.2e-125;
Matches 268; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 1 MHKAGLLGLCARAWNSVYRMASGMTRDPLANKVALVTASTDGIGFAIRLAQDAHV 60
Db 1 mhmarllglcawarksvrmassrmtrdpltnkvalvtastdgigfaiarraqdahv 60
QY 61 VSSRKQNVDOAVATLOGEGLSVTGTVCHVCKAEDRELRLVATVAVKLHGIDILVSNAAVN 120
Db 61 vssrkqnvddqavatlqgeglsvtgtvchvgkaedrelrlvatavklhgidilvsnaavn 120
QY 121 PFFGSIIMDVTEEVWDKTLIDINVKAPALMTKAVVPEMEKRGGSVVIVSSIAAFSPSPGFS 180
Db 121 pffgsimdvteevwkdltidinvkapalmckavvpemekrggsvvsvssiaafspspgfs 180

QY 181 PYNVSKTALLGLTKTALIELAPNIRVNCIAPGLIKTSFSLMLWMDKEESMKETLIR 240
 Db 181 pynvsktallglntlaelapnirvnciapglktsfslmlwmdkeesmketlir 240
 QY 241 RLGEPEDCAGIVFLCSEADASYITGETWVGGGTPSRL 278
 Db 241 rlgepedcagivflcsedasyitgetvvvggtpsl 278

RESULT 4
 AAB58822
 ID AAB58822 standard; Protein; 183 AA.
 AC AAB58822;
 XX
 DT 27-MAR-2001 (first entry)
 DE Breast and ovarian cancer associated antigen protein sequence SEQ ID 530.
 DE Human; breast cancer; ovarian cancer; cytostatic; immunosuppressive;
 KW neurotropic; neuroprotective; antiviral; antiallergic; hepatotropic;
 KW antidiabetic; antiinflammatory; antiulcer; vulnery; anticonvulsant;
 KW antibacterial; antifungal; antiparasitic; cardiac; immune disorder;
 KW Addison's disease; allergy; autoimmune haemolytic anaemia;
 KW autoimmune thyroiditis; diabetes mellitus; Crohn's disease;
 KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis;
 KW cardiovascular disorder; wound healing; neurological disease.
 XX Homo sapiens.
 OS
 XX
 XX
 PN WO20005173-A1.
 PD 21-SEP-2000.
 XX
 XX 08-MAR-2000; 2000WO-0505881.
 XX 12-MAR-1999; 99US-012470.
 XX
 PA (HUMA-) HUMAN GENOME SCI. INC.
 PI Rosen CA, Ruben SM;
 DR WPI; 2000-611515/58.
 DR N-PSDB; AAF21725.
 XX
 XX New human breast and ovarian cancer associated gene sequences and the
 PT polypeptides encoded by these genes, useful in the prevention,
 PT treatment and diagnosis of cancer, immune disorders, cardiovascular
 PT disorders and neurological diseases.
 XX
 PS Claim 11; Page 963-964; 1299pp; English.
 PS
 XX Sequences AAF21614 - AAF22031 represent DNA sequences encoding human
 CC proteins AAB58711 - AAB59128. The DNA and protein sequences are
 CC associated with breast and ovarian cancer. Included in the invention are
 CC sequences AAF22032 - AAF22040 and AAB59129 which are used in the
 CC isolation and characterisation of the DNA and protein sequences of the
 CC invention. The breast and ovarian cancer associated DNA, protein, agonist
 CC or antagonist sequences exhibit cytostatic; immunosuppressive;
 CC neurotropic; neuroprotective; antiviral; antiallergic; hepatotropic;
 CC antidiabetic; antiinflammatory; antiulcer; vulnery; anticonvulsant;
 CC antibacterial; antifungal; antiparasitic and cardiac activity. The
 CC polynucleotide and protein sequences are used in the diagnosis of cancer,
 CC particularly breast and ovarian cancer. The nucleic acid sequences,
 CC proteins, agonists and antagonists may also be used in the diagnosis,
 CC prevention and treatment of immune disorders e.g. Addison's disease,
 CC allergies, autoimmune haemolytic anaemia, autoimmune thyroiditis,
 CC diabetes mellitus, Crohn's disease, multiple sclerosis, rheumatoid
 CC arthritis and ulcerative colitis; cardiovascular disorders such as
 CC myocardial ischaemia; wound healing; neurological diseases such as
 CC cerebral anoxia and epilepsy; and infectious diseases.
 XX
 SQ Sequence 183 AA;

Query Match 46.4%; Score 649.5; DB 21; Length 183;
 Best Local Similarity 55.2%; Pred. No. 4e-57;
 Matches 148; Conservative 2; Mismatches 21; Indels 97; Gaps 3;
 QY 16 SVRWASSGTRRDRPLANKVALVTASTDGIGFATARRLAODGAHVSVSRKQONVDQAV-- 73
 Db 8 svrmassmtrrdpltnkvalvtastdgigf-----aspvvprrtgprgrgpeaaccg 61
 QY 74 ---ATLQGEGLSVTGVCHVKAEDRELRVATVKLHGDIILVSNAAVNPFFGCSIMDVT 130
 Db 62 pgggtlqgeglsvtgcxxxgkaedrervlrvatavklhgdiilvsnaaavnpffgsimdvt 121
 QY 131 EYVDKTLINVKAPALMTKAVVPEMEKRGGSVVIVSSIAAFSPSPGSPYNNVSKTALL 190
 Db 122 evwdk----- 127
 QY 191 GLTKTALIELAPNIRVNCIAPGLIKTSFSLMLWMDKEESMKETLIRRLGPEDCAG 250
 Db 128 -----lwm dkeesmketlirrlgpedcag 155
 QY 251 IVSFLCSEADASYITGETWVGGGTPSRL 278
 Db 156 ivsflcsedasyitgetvvvggtpsl 183

RESULT 5
 AAG41219
 ID AAG41219 standard; Protein; 254 AA.
 AC AAG41219;
 DT 18-OCT-2000 (first entry)
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 51256.
 KW Protein identification; signal transduction pathway; metabolic pathway;
 KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence.
 XX Arabidopsis thaliana.
 XX EP1033405-A2.
 PD 06-SEP-2000.
 XX 25-FEB-2000; 2000EP-0301439.
 XX 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 28-APR-1999; 99US-0130891.
 PR 30-APR-1999; 99US-0132048.
 PR 30-APR-1999; 99US-0132407.
 PR 04-MAY-1999; 99US-0132484.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 07-MAY-1999; 99US-0132487.
 PR 11-MAY-1999; 99US-0132863.
 PR 14-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.

PA	(DAIL) DAICEL CHEM IND LTD.
XX	
XX	Yamamoto H;
XX	
XX	WPI: 2000-118183/11.
DR	N-PSDB; AA245749.
DR	
XX	Producing optically pure (S)-4-halo-3-hydroxybutyric acid ester -
PT	
XX	
XX	Claim 6; Page 19-20; 34pp; English.
PS	
XX	
CC	The present sequence represents a beta-ketoacyl-ACP reductase protein
CC	of bacillus subtilis. The beta-ketoacyl-ACP reductase enzyme constitutes
CC	a Type II fatty acid synthetase. The enzyme has an extremely high
CC	reducing activity and stereoselectivity towards 4-chloroacetoacetic
CC	acid ester. The specification describes a method for producing a
CC	(S)-4-halo-3-hydroxybutyric acid ester. The method comprises
CC	asymmetrically reducing 4-halo-acetoacetic acid ester or its
CC	derivative with beta-ketoacyl-acyl carrier protein reductase
CC	constituting Type II fatty acid synthase, or acetoacetyl-CoA
CC	reductase constituting the polybeta-hydroxy fatty acid biosynthesis
CC	system. The novel method is used to produce optically active
CC	4-halo-3-hydroxybutyric acid ester, with a high purity.
XX	
XX	Sequence 248 AA;

Query Match 28.9%; Score 404; DB 21; Length 248;
Best Local Similarity 37.7%; Pred. No. 2.6e-32;
Matches 93; Conservative 50; Mismatches 100; Indels 4; Gaps 3

Qy	28	DPLANKVALVTA	TDGIGFAIARR	LAQGAHVVS-SRKQONVDQAVATLQEGCLSVTGT	86
Db	2	dmlndktaivt	gasrgisr	atalaksganvvnysgneakaneevdeiksmgrkaiav	61
Qy	87	VCHVGKAEDRE	RLVATAKVLHCG	TDILVSNAAVNPFGSTMDVTEFYWDKTLDINWKAPA	146
Db	62	kadvspedvq	miiketisv	fstidilvnnagitr-dnlmrmkedewdvvinhklgvf	120
Qy	147	LMTKAVVPEMK	RGGGSVVIVSS	TAAFSPPGFSPYNVSKTALGLTLAIELAPNIR	206
Db	121	nctkavtrgm	ukqrgri	ainvssivvgsgnpgqnyvaakgvgilckssakelasrnit	180
Qy	207	VNCLAPGLLKT	SFRMLNMDKE	BESMKETULIRRLCEPCAGIVSFLCSEDAASYITGE	266
Db	181	vnaia	pgfstidmtdkl--akdvqeml	kqiplarfgepsdvssvflasgarmtcg	238
Qy	267	TVVVG	GG		273
Db	239	tlh	dq		245

QY	207	VNCLAPGLIKTSF	SRMLNMDKEES	MRKTIIRRLG	PEPCAGIVSFLC	SEDASYITGE	266
		: :	:	: :	: :	: :	
Db	181	vnaiapgfistd	mtckl--akdv	gemikqiplarf	gepsdvsvvtf	lasegarymtg	238
		: :	:	: :	: :	: :	
QY	267	TVVVGGS	273				
		: :	:	: :	: :	: :	
Db	239	tlhldgg	245				
		: :	:	: :	: :	: :	

RESULT	9	
AAG02114		
ID	AAG02114	standard; Protein; 79 AA.
XX		
XX		
AC	AAG02114;	
XX		
XX		
DT	06-OCT-2000	(first entry)
DE	Human secreted protein, SEQ ID NO: 6195.	

DE	Human secreted protein, SEQ ID NO: 6195.	
XX		
XX		
XX	Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;	
KW	gene therapy; Chromosome mapping.	
XX		
XX		
OS	Homo sapiens.	
XX		
PN	EP1033401-A2.	
XX		
PD	06-SEP-2000.	
XX		
XX	21-FEB-2000; 2000EP-0200610.	
XX		

26-FEB-1999; 99US-0122487.
(GEST) GENSET.
Dumas Milne Edwards J, Duclert A, Giordano J;
WPI; 2000-500381/45.
N-PSDB; AAC02120.
New nucleic acid that is a 5' expressed sequence tag (5' EST) for
obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
diagnostic, forensic, gene therapy and chromosome mapping procedures -
Claim 13; SEQ ID 6195; 71pp + CD-ROM; English.
The present sequence is a polypeptide encoded by one of a large number
of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
were prepared from total human RNAs or polyA+ RNAs derived from 30
different tissues. EST sequences usually correspond mainly to the 3',
untranslated region (UTR) of the mRNA because they are often obtained
from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
isolating cDNA sequences derived from the 5' ends of mRNAs and even in
those cases where longer cDNA sequences have been obtained, the full 5'
UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
ends and can therefore be used to obtain full length cDNAs and genomic
DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
chromosome mapping procedures. They are used to obtain upstream
regulatory sequences and to design expression and secretion vectors.
Sequence 79 AA;
Query Match 26.4%; Score 370; DB 21; Length 79;
Best Local Similarity 96.2%; Pred. NO. 1.3e-29;
Matches 76; Conservative 0; Mismatches 3; Indels 0; Gaps
QY 1 MHRAGLLGLCARAWNSVRMASSGWRDPLANKVALVTASTDIGFAIARRLAQDGAHV 60
Db 1 mxxaxllglccarawnsvrmassgwrdrplankvalvtastdigfaiarlqdgahv 60
QY 61 VSSRKQNVQDAVATLQGE 79
Db 61 vssrkqnvqavatlqge 79
RESULT 10
AAG31818
ID AAG31818 standard; Protein; 262 AA.
XX AC AAG31818;
XX AC
XX AC
17-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 38275.
DE Arabidopsis thaliana
XX Protein identification; signal transduction pathway; metabolic pathway;
XX hybridisation assay; genetic mapping; gene expression control; promoter
XX termination sequence.
OS Arabidopsis thaliana.
XX Arabidopsis thaliana.
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 99US-0121825.
XX 05-MAR-1999; 99US-0123180.
XX 09-MAR-1999; 99US-0123548.
XX 23-MAR-1999; 99US-0125788.
XX 25-MAR-1999; 99US-0126264.
XX 29-MAR-1999; 99US-0126785.

PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.

Query Match 25.2%; Score 353; DB 21; Length 262;
Best Local Similarity 35.0%; Pred. No. 3.7e-27;
Matches 89; Conservative 47; Mismatches 110; Indels 8; Gaps

QY	24	MTRRDPLANKVALVTASTGIGFAIARRLAODGAHVHVSRRKQNVDDQAVATLQGBGLSV	83
DB		:	
Db	1	mkrkwskgmtalvtgagsgiyvaiveelagfgarihvcdiseaklngalsewekkqfgy	60
QY	84	TGVCHVHGKAEDRERIVAN-AVKLHGIDILYSNAAV---NPFFGSIMDVTEEVWDKTLD	139
Db		:	
Db	61	sgsvdcvvarperceelmqtvsqfdknllvsnvgvirskp----tteyteddfafhis	116
QY	140	INVKAPALMTKAVVPMEKRGGSGVVIVSSIAAFSPGFSFYNVSKTALLGLTKTLAIE	199
Db			
Db	117	snveaaayhfisqlshpllkasgygsiifvssiaqvtsfdagsiygtktgaliglaknlace	176
QY	200	LAPNRIRVNCLAPGLIKTSFRSMLWMMDKEKESMKETLRRLGPEDCAGIVSLCSED	259
Db			
Db	177	wakdgrananapnvnltplsdyledsfkallstrtlprgvgepnvasivaflicipa	236
QY	260	ASYITGETVVVGCG 273	
Db			
Db	237	asyitgtticvdgg 250	

RESULT 11
AAG92082
ID AAG92082 standard; Protein; 292 AA.
XX AC
AC AAG92082;
DT DT
XT XT
DE DE
XX C glutamicum protein fragment SEQ ID NO: 5836.
KW Coryneform bacterium; amino acid synthesis; vitamin; saccharide;
KK organic acid synthesis.
XX Corynebacterium glutamicum.
XX EP1108790-A2.
PN PD
XX 20-JUN-2001.

CC containing them are used to map genomes of organisms related to
 CC C. glutamicum, identify and localise C. glutamicum sequences of interest,
 CC in evolutionary studies, in determining SMP protein regions required
 CC for function, in modulating SMP protein activity, in modulating the
 CC metabolism of sugars, and in modulating high-energy molecule production
 CC in a cell (i.e. ATP, NADPH).
 XX
 SQ Sequence 295 AA;

Query Match 24.9%; Score 349; DB 22; Length 295;
 Best Local Similarity 34.5%; Pred. No. 1.1e-26;
 Matches 86; Conservative 45; Mismatches 114; Indels 4; Gaps 2;

QY 30 LANKVALVTASTDGIQFAIARRLAQDGAHVVS--SRKQNVNDAVATLQEGSLSVGTGV 87
 Db 48 lkrkalitgdsigaavalaayaregadvaiaylpeeqadadrvlqaeetgkafsf 107
 QY 88 CHVGKAEDERLVATAVKLHGIDILVSNAAVNPFFGSIMDVTEVMDKTLIDINVKAPAL 147
 Db 108 gdlrdpeycrslvqetnalggdlvnnasrqwvapgltedntdqtqnlvysfr 167
 QY 148 MTKAVVPEMEKRGGSVVIVSSIAAFSPGFSFYNVSKTALLGLTKTLATELAPRIRV 207
 Db 168 vtkaaphl--kpgssiftssiqaypsetlldyamtkaalnnlskglasllgdgivr 225
 QY 208 NCLAPGLIKTSFMRMLNDKEESMKETLIRLGPEDCAGIVSFICSDASYITGET 267
 Db 226 nsvapgpwtlqpsghgqkqiegfgqhaghpvelagayvflasdeasyvvyvget 285
 QY 268 VVGGGTPS 276
 Db 286 lgvctgtpt 294

RESULT 13
 AAB79401
 ID AAB79401 standard; Protein: 295 AA.
 XX
 AC AAB79401;
 XX
 DT 30-Apr-2001 (first entry)
 XX
 DE Corynebacterium glutamicum SMP protein sequence SEQ ID NO:318.
 XX
 KW Corynebacterium glutamicum; carbon metabolism and energy production;
 KW SMP protein; sugar metabolism and oxidative phosphorylation protein;
 KW fine chemical production; organic acid; proteinogenic amino acid;
 KW nonproteinogenic amino acid; purine base; pyrimidine base; nucleoside;
 KW nucleotide; lipid; saturated fatty acid; unsaturated fatty acid; diol;
 KW carbohydrate; aromatic compound; vitamin; cofactor; polyketide; enzyme;
 KW diagnosis; Corynebacterium diphtheriae; evolutionary study.
 XX
 OS Corynebacterium glutamicum.
 XX
 PN W0200100844-A2.
 XX
 PD 04-JAN-2001.
 XX
 PF 23-JUN-2000; 2000WO-IB00943.
 XX
 PR 25-JUN-1999; 99US-0141031.
 PR 08-JUL-1999; 99DE-1031412.
 PR 08-JUL-1999; 99DE-1031413.
 PR 08-JUL-1999; 99DE-1031419.
 PR 08-JUL-1999; 99DE-1031420.
 PR 08-JUL-1999; 99DE-1031424.
 PR 08-JUL-1999; 99DE-1031428.
 PR 08-JUL-1999; 99DE-1031431.
 PR 08-JUL-1999; 99DE-1031433.
 PR 08-JUL-1999; 99DE-1031434.
 PR 08-JUL-1999; 99DE-1031510.
 PR 08-JUL-1999; 99DE-1031562.

PR 08-JUL-1999; 99DE-1031634.
 PR 09-JUL-1999; 99DE-1032180.
 PR 09-JUL-1999; 99DE-1032227.
 PR 09-JUL-1999; 99DE-1032230.
 PR 09-JUL-1999; 99US-0143208.
 PR 14-JUL-1999; 99DE-1032924.
 PR 14-JUL-1999; 99DE-1032973.
 PR 14-JUL-1999; 99DE-1033005.
 PR 27-AUG-1999; 99DE-1040765.
 PR 31-AUG-1999; 99US-0151572.
 PR 03-SEP-1999; 99DE-1042076.
 PR 03-SEP-1999; 99DE-1042079.
 PR 03-SEP-1999; 99DE-1042086.
 PR 03-SEP-1999; 99DE-1042087.
 PR 03-SEP-1999; 99DE-1042088.
 PR 03-SEP-1999; 99DE-1042095.
 PR 03-SEP-1999; 99DE-1042123.
 PR 03-SEP-1999; 99DE-1042125.
 XX
 PA (BADI) BASF AG.
 XX
 PI Pompejus M, Kroeger B, Schroeder H, Zeider O, Haberhauer G;
 XX
 DR WPI: 2001-061975/07.
 DR N-PSDB; AAF71518.

New isolated Corynebacterium glutamicum nucleic acid encoding a sugar
 metabolism and oxidative phosphorylation protein for production or
 modulation of production of fine chemicals e.g. amino acids,
 carbohydrates or enzymes

Claim 20; Page 587-588; 1246pp; English.

AAF71360 to AAF71750 encode the Corynebacterium glutamicum sugar
 metabolism and oxidative phosphorylation (SMP) proteins given in
 AAB79243 to AAB 79633 which are involved in carbon metabolism and
 energy production. The C. glutamicum SMP gene can be used in vectors
 (II) for expression in host cells and production or modulation of
 production of fine chemicals, such as, an organic acid, a proteinogenic
 or nonproteinogenic amino acid (preferred), a purine or pyrimidine base,
 a nucleoside, a nucleotide, a lipid, a saturated or unsaturated fatty
 acid, a diol, a carbohydrate, an aromatic compound, a vitamin, a
 cofactor, a polyketide, or an enzyme. The presence of (I) or SMP proteins
 (III) encoded by them are used for diagnosing the presence or activity of
 Corynebacterium diphtheriae in a subject. (I), (II), (III) or host cells
 containing them are used to map genomes of organisms related to
 C. glutamicum, identify and localise C. glutamicum sequences of interest,
 in evolutionary studies, in determining SMP protein regions required
 for function, in modulating SMP protein activity, in modulating the
 metabolism of sugars, and in modulating high-energy molecule production
 in a cell (i.e. ATP, NADPH).

Sequence 295 AA;

Query Match 24.9%; Score 349; DB 22; Length 295;
 Best Local Similarity 34.5%; Pred. No. 1.1e-26;
 Matches 86; Conservative 45; Mismatches 114; Indels 4; Gaps 2;

QY 30 LANKVALVTASTDGIQFAIARRLAQDGAHVVS--SRKQNVNDAVATLQEGSLSVGTGV 87
 Db 48 lkrkalitgdsigaavalaayaregadvaiaylpeeqadadrvlqaeetgkafsf 107
 QY 88 CHVGKAEDERLVATAVKLHGIDILVSNAAVNPFFGSIMDVTEVMDKTLIDINVKAPAL 147
 Db 108 gdlrdpeycrslvqetnalggdlvnnasrqwvapgltedntdqtqnlvysfr 167
 QY 148 MTKAVVPEMEKRGGSVVIVSSIAAFSPGFSFYNVSKTALLGLTKTLATELAPRIRV 207
 Db 168 vtkaaphl--kpgssiftssiqaypsetlldyamtkaalnnlskglasllgdgivr 225
 QY 208 NCLAPGLIKTSFMRMLNDKEESMKETLIRLGPEDCAGIVSFICSDASYITGET 267
 Db 226 nsvapgpwtlqpsghgqkqiegfgqhaghpvelagayvflasdeasyvvyvget 285

	QY	92	KAEEDRLRVATAVKLHGIDILVSNAAVNPPFGSIMDVTEEVMKDTLDINVKAPALMTKA	151
	Dd		I:: :: : I : : : I : : : : : :	
	Dd	65	dadevkamikevvsgfgslvdvlvnagitr-dnlImrkgewdvidunlkgvfncigk	123
	QY	152	VPEMEKRGGGSVIVSSIAAFSPGFSPYNVSXTALLGLTKTLAIELAPRNRVNCLA	211
	Dd	124	atpqmlrqsгалinlvssvgavpgqganyvatkvlgltksaaarelascrgitcnava	183
	QY	212	PGLIKTSFRMLWMDKEESMKETLRIRLCEPEDCAGISVFCLSEDSASITGTVVVG	271
	Dd	184	pbfivsdmtdal--sdeikeymgtlpiarfogtdiantvaflasdkakyitqtihvn	241
	QY	272	GG 273	
	Dd	242	gg 243	
		RESULT 15		
	AAC	AAE02195		
	ID	AAE02195 standard; Protein; 246 AA.		
	XX	AC	AAE02195;	
	DD			
	DT	31-JUL-2001 (first entry)		
	XX	S. aureus NADPH-dependent beta-ketoacyl-ACP reductase (FabG).		
	DE	FabG; high throughput method; fatty acid biosynthesis; therapy;		
	KW	bacterial enzyme; biological agent screening; otitis media; empyema;		
	KW	bacterial tracheitis; acute epiglottitis; thyroiditis; lung abscess;		
	KW	infective endocarditis; secretory diarrhoea; splenic retroperitoneal;		
	KW	intrarenal; perinephric; cerebral; cutaneous; abscess; blepharitis;		
	KW	conjunctivitis; keratitis; endophthalmitis; preseptal; orbital;		
	KW	cellulitis; dacryocystitis; epididymitis; abscess; toxic shock syndro-		
	KW	mepigo; folliculitis; wound infection; bacterial myositis;		
	KW	septic arthritis; osteomyelitis; beta-ketoacyl-ACP reductase;		
	KW	acyl carrier protein.		
	XX			
	OS	Staphylococcus aureus 'WCUH 29'.		
	XX			
	PN	WO200130988-A1.		
	XX			
	PD	03-MAY-2001.		
	XX			
	PF	26-OCT-2000; 2000WO-US29451.		
	XX			
	PR	27-OCT-1999; 99US-0161775.		
	XX	(SMTK) SMITHKLNE BEECHAM CORP.		
	PA	(SMIK) SMITHKLNE BEECHAM PLC.		
	PI	Dewolf W, Kallender H, Lonsdale JT;		
	XN			
	DR	WIPI ; 2001-316332/33.		
	DR	N-PSDB; AAD06206.		
	PT	High throughput method for screening for biological agents against		
	PT	fatty acid biosynthesis comprises contacting a bacterial enzymatic		
	PT	pathway with enzymes e.g. malonyl-CoA:ACP transacylase -		
	XS	Claim 1; Page 10; 94pp; English.		

The present invention relates to a high throughput method for screening biological agents affecting fatty acid biosynthesis, comprises contacting a bacterial enzymatic pathway with enzymes. The method is used for screening biological agents affecting fatty acid biosynthesis. Agonists and antagonists of fab (fatty acid biosynthesis) are used to inhibit, prevent or treat diseases such as infections of the upper respiratory tract (e.g. otitis media, bacterial tracheitis, acute epiglottitis, thyroiditis), lower respiratory (e.g. emphysema, lung abscess), cardiac (e.g. infective endocarditis), gastrointestinal (e.g. secretory diarrhoea, splenic abscess, retroperitoneal abscess), central

Sequence 246 AA:

3.

92 KAEDRERLVATAVKLHGGIDILVSNAAVNPFEGSIMDVTEEVWDKTLDINVKAPALMTKA-151

65 dadevkamikevvsqfsgldvlnnagitr-dnllmrmkegewddvidtnlkgvfnciqk 123

152 VVPEMERGGGSWIVSSIAAFSPSPGFSPYNVSKTALLGLTKTLAIELAPRNIRVNCLA 211

[illegible][illegible]

0
E
C
4
4
C
r
C

—

Search completed: January 22, 2002, 15:19:05
Job time: 195 sec

GenCore version 4.5
Copyright (c) 1993-2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2002, 15:17:30 ; Search time 12.54 Seconds
(without alignments)
(498.877 Million cell updates/sec)

Title: US-09-866-034-2

Perfect score: 1399

Sequence: 1 MHKAGLLGICARAWNSVRMA.....DASYITGETVVGCGTPSRL 278

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match: 0%

Maximum Match: 100%

Listing first 45 summaries

Database : Issued Patents_AA:*

- 1: /cgn2_6/ptodata/2/1aa/5A_COMB.pap:*
- 2: /cgn2_6/ptodata/2/1aa/5B_COMB.pap:*
- 3: /cgn2_6/ptodata/2/1aa/6A_COMB.pap:*
- 4: /cgn2_6/ptodata/2/1aa/6B_COMB.pap:*
- 5: /cgn2_6/ptodata/2/1aa/PT05_COMB.pap:*
- 6: /cgn2_6/ptodata/2/1aa/backfiles1.pap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	348	24.9	246	3	US-09-238-481-2
2	335.5	24.0	313	4	US-09-413-814-9
3	328.5	23.5	345	3	US-08-793-035-9
4	328.5	23.5	315	3	US-08-793-035-10
5	324.5	23.2	333	1	US-08-440-856A-4
6	324	23.2	262	4	US-09-363-189B-6
7	321	22.9	247	1	US-08-241-766-13
8	313	22.4	244	1	US-08-375-962B-13
9	313	22.4	244	2	US-08-562-114B-13
10	313	22.4	244	4	US-08-729-594A-13
11	312.5	22.3	256	1	US-08-594-808B-7
12	306.5	21.9	337	1	US-08-440-856A-3
13	299.5	21.4	303	3	US-09-002-298-1
14	296.5	21.2	257	4	US-09-287-097-2
15	295	21.1	244	1	US-08-762-129-1
16	295	21.1	244	2	US-09-090-567-2
17	295	21.0	271	2	US-07-637-865-2
18	294	21.0	244	1	US-08-762-129-3
19	279	19.9	246	6	5229279-7
20	272.5	19.5	255	4	US-08-815-225-4
21	270	19.3	251	3	US-08-822-322-9
22	270	19.3	251	4	US-09-466-109-9
23	270	19.3	252	3	US-08-822-322-8
24	270	19.3	252	4	US-09-466-109-8
25	266	19.0	244	1	US-08-762-129-4
26	258.5	18.5	283	4	US-09-367-012-1
27	253	18.1	273	6	5512669-4

28 248 17.7 359 1 US-08-440-856A-8 Sequence 8, Appli
29 246.5 17.6 248 4 US-09-385-028-11 Sequence 11, Appli
30 241 17.2 335 3 US-09-002-298-7 Sequence 7, Appli
31 237 16.9 295 3 US-09-002-298-5 Sequence 5, Appli
32 230.5 16.5 292 3 US-09-109-205-2 Sequence 2, Appli
33 226 16.2 335 3 US-09-002-298-6 Patent No. 5229279
34 224 16.0 263 6 5229279-4 Sequence 19, Appli
35 221 15.8 335 3 US-09-109-205-19 Sequence 7, Appli
36 217.5 15.5 231 2 US-08-336-198C-7 Sequence 2, Appli
37 217.5 15.5 295 4 US-09-026-482B-2 Sequence 12, Appli
38 197.5 14.1 327 1 US-08-375-962B-12 Sequence 12, Appli
39 197.5 14.1 327 2 US-08-562-114B-12 Sequence 12, Appli
40 197.5 14.1 327 4 US-08-729-594A-12 Sequence 12, Appli
41 191.5 13.7 261 4 US-08-815-225-2 Sequence 2, Appli
42 191.5 13.7 261 4 US-08-815-225-3 Sequence 3, Appli
43 183.5 13.1 388 4 US-08-980-832-41 Sequence 41, Appli
44 173 12.4 325 3 US-08-581-148C-4 Sequence 4, Appli
45 170.5 12.2 309 3 US-09-109-205-1 Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-238-481-2
; Sequence 2, Application US/09238481
; Patent No. 6110704
; GENERAL INFORMATION:
; APPLICANT: Huang, Jianzhong
; APPLICANT: McDevitt, Damien
; TITLE OF INVENTION: Fabg
; FILE REFERENCE: GM10192
; CURRENT APPLICATION NUMBER: US/09/238, 481
; CURRENT FILING DATE: 1999-01-28
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 246
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-238-481-2

Query Match 24.9%; Score 348; DB 3; Length 246;
Best Local Similarity 32.6%; Pred. No. 3.8e-30;
Matches 79; Conservative 57; Mismatches 102; Indels 4; Gaps 3;
QY 33 KVALVTASTDGGFAIARRLAODGAHVVS-SRKOQNVDAVATLQEGLSVTGTVCHVG 91
DB 5 KSAIVTASRGIGRSIALQLAEEGVNAVYAGSKEKAEAVVEEIKAKGVDSFAIQANVA 64
QY 92 KAEDRERLVATAVKLHGIDILVSNRAVNPFFGSIIMDVTEEVWDKTLDINVKAPALMTKA 151
DB 65 DADVKAMKEVVSQFSLDLVNNAGITR-DNLLMRKKEQEMDDVDIDNLSKGVFNICOK 123
QY 152 VVPEMEKGGGVSVISSTAAPSPGSPYVNTALLGLTKTLAIELAPNRVNCIA 211
DB 124 ATPQMLRQSGAIINLSVVGAVGNPQAVYATKAGVIGLTSAARELASRGITVNAVA 183
QY 212 PCLIKTSFRMLMMDKEESMKETLRRRLRCEPCAGIVSFLCSEDASYITGETVVVG 271
DB 184 PGFIVSDMTDAL--SDELKEQMLTOIPLARFGQDITANTVAFLASDKAKYITGOTIHVN 241
QY 272 GG 273
DB 242 GG 243

RESULT 2
US-09-413-814-9
; Sequence 9, Application US/09413814
; Patent No. 6225064
; GENERAL INFORMATION:

APPLICANT: Gesellschaft für Biotechnologische Forschung mbH
APPLICANT: Bristol-Myers Squibb, Co.
APPLICANT: Beyer, Stefan
APPLICANT: Bloecker, Helmut
APPLICANT: Brandt, Petra
APPLICANT: Cino, Paul M
APPLICANT: Dougherty, Brian A
APPLICANT: Goldberg, Steven L
APPLICANT: Hofle, Gerhard
APPLICANT: Mueller, Joachim
APPLICANT: Reichenbach, Hans
TITLE OF INVENTION: DNA sequences for enzymatic synthesis of polyketide or
FILE REFERENCE: PCT/US 99/23535
CURRENT APPLICATION NUMBER: US/09/413,814
CURRENT FILING DATE: 1999-10-07
EARLIER APPLICATION NUMBER: DE 198 46 493.2
EARLIER FILING DATE: 1998-10-09
NUMBER OF SEQ ID NOS: 107
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 9
LENGTH: 313
TYPE: PRT
ORGANISM: Sorangium cellulosum
US-09-413-814-9

Query Match 24.0%; Score 335.5; DB 4; Length 313;
Best Local Similarity 33.6%; Pred. No. 1.3e-28;
Matches 83; Conservative 50; Mismatches 105; Indels 9; Gaps 4;

QY 30 LANKVALVTASDTGTFATRLAQDGAHVYVSSRRKQNVDOAVATLOGEGLSVGTGVC 88
DB 50 LAGRALVTGSSRGIGKATLRLAEOGADVAVNHSNKDAAEQTAETRLGRRIMVYQA 109
QY 89 HVGKAEDRRLVATVKLHGGIDILVSNAAVNPFFGSDVTEEVWDTLIDINVKAPALM 148
DB 110 DVTRNAAELFSSVEAQIGPDIDHVNNGV-DFFPKPLAAMTDDDRNVMSNLSVHYL 168
QY 149 KAVVPEMEKRGSGSVIVYSSIAAFSPGSPFPYNSKTLGLTKTLATELAPRNR 206
DB 169 CRAAVARMQRSGRIINIGLSPFAIRCAPNVAYSTAKTGVILTRSLATEEAPHGIL 228
QY 207 VNCLAPGLIKTSFSEMLMWDKKEESMKETLRLRGLPEDCAGIVSLFCSDESAYITGE 266
DB 229 VNCVSPGLIDNGY-----LPPAQKEWMERRVPNGRLGRASEVADAVAFIADRSYVSGA 283
QY 267 TVVGGG 273
DB 284 NIAVAGG 290

RESULT 3
US-09-793-035-9
Sequence 9, Application US/08/793035
Patent No. 6011201
GENERAL INFORMATION:
APPLICANT: Slabas, Antoni R.
APPLICANT: White, Andrew
APPLICANT: Chase, Dianne
APPLICANT: Elborough, Keiran
APPLICANT: Fentem, Phillip A.
TITLE OF INVENTION: B-ketoacyl ACP Reductase Genes From
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: US
ZIP: 77210-4433
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/793,035
FILING DATE: 28-JUL-1997
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9414622.2
FILING DATE: 20-JUL-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB95/01678
FILING DATE: 17-JUL-1995
ATTORNEY/AGENT INFORMATION:
NAME: Kammerer, Patricia A.
REGISTRATION NUMBER: 29,775
REFERENCE/DOCKET NUMBER: MOBT:132
TELECOMMUNICATION INFORMATION:
TELEPHONE: 713.787.1400
TELEFAX: 713.787.1440
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 315 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-793-035-9

Query Match 23.5%; Score 328.5; DB 3; Length 315;
Best Local Similarity 33.1%; Pred. No. 7.9e-28;
Matches 81; Conservative 53; Mismatches 100; Indels 11; Gaps 5;

QY 34 VALVTASTDGTGTFATRLAQDGAHVYV-----SSRQNVDOAVATLOGEGLSVGTGVC 89
DB 74 VVVVTGASRGIGKATLRLAEOGADVAVNHSNKDAAEQTAETRLGRRIMVYQA 130
QY 90 VGKAEDRRLVATVKLHGGIDILVSNAAVNPFFGSDVTEEVWDTLIDINVKAPALM 149
DB 131 VSKEADVEMAMKTAIDANGTIDVVVNNAGITR-DTLLIRMKKSQWDEVIDLNLTVFLCT 189
QY 150 KAVVPEMEKRGSGSVIVYSSIAAFSPGSPFPYNSKTLGLTKTLATELAPRNRVNC 209
DB 190 QAAATKIMMKRRGRIINIASVVGGLIGNTQAAKAGVIGFSKTAAREGASRNINNV 249
QY 210 LAPGLIKTSFSEMLMWDKKEESMKETLRLRGLPEDCAGIVSLFCSDESAYITGETV 268
DB 250 VCPGFIASDMTAKLGEDMEKK--ILGTIPLGRYGQPEDVAGLVEFLALSPAASYITGOAF 307
QY 269 VVGGG 273
DB 308 TIDGG 312

RESULT 4
US-08-793-035-10
Sequence 10, Application US/08/793035
Patent No. 6011201
GENERAL INFORMATION:
APPLICANT: Slabas, Antoni R.
APPLICANT: White, Andrew
APPLICANT: Chase, Dianne
APPLICANT: Elborough, Keiran
APPLICANT: Fentem, Phillip A.
TITLE OF INVENTION: B-ketoacyl ACP Reductase Genes From
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: TX

```

, ZIP: 20037
,
, COMPUTER READABLE FORM:
,
, MEDIUM TYPE: Floppy disk
,
, COMPUTER: IBM PC compatible
,
, OPERATING SYSTEM: PC-DOS/MS-DOS
,
, SOFTWARE: PatentIn Release #1.0, Version #1.25
,
, CURRENT APPLICATION DATA:
,
, APPLICATION NUMBER: US/08/440.856A
,
, FILING DATE: 15-MAY-1995
,
, CLASSIFICATION: 800
,
, ATTORNEY/AGENT INFORMATION:
,
, NAME: MILLMAN, ROBERT A.
,
, REGISTRATION NUMBER: 36,217
,
, REFERENCE/DOCKET NUMBER: 05463-20001.00
,
, TELECOMMUNICATION INFORMATION:
,
, TELEPHONE: (202) 887-1517
,
, TELEFAX: (202) 887-0763
,
, TELEX: 706141
,
, INFORMATION FOR SEQ ID NO: 4:
,
, SEQUENCE CHARACTERISTICS:
,
, LENGTH: 333 amino acids
,
, TYPE: amino acid
,
, STRANDEDNESS: single
,
, TOPOLOGY: linear
,
, US-08-440-856A-4

```

Best local similarity	52.5%	Freq.	NO.	2.4e	27,
Matches	102;	Conservative	37;	Mismatches	108;
Indels	63;	Gaps	11;		

Qy		13	ANNSVRMASSGTRRDPLANKVALVATSTDGTGFIAIRRLAQDGHHVVSVSRKQNWDQA	72
Db		37	AWDSPNGAPTMPKR---LEGKVAIYTGARGIGEAIVLEFKHGAKVIA-----DIDDA	89
Qy		73	VATLOGEGLSV-----TGTV-CHVGKAEDREBLVATAVLHCGTDILYSNAVNPFPG--	124
Db		90	A----GEALAAALPHGVFVRCDSVEEDVERAVERAVARYGRDLVLCNNAGV---LGRQ	142
Qy		125	-----SIMDVTEEVMDKTLDINVKPALMTKAVPEMEKRGGSVVIVSSIIAASPSPGF	179
Db		143	TAAKSIIISFDAGEFDRVLRVNLAAGAALGMKHAALAMTORRAGSIISVASVAGVLGGLP	202
Qy		180	SPYNYSKTALLGITLTATELAPRIRNVCNLAGLIKTSFSMLMM-----	225
Db		203	HAYTASKHAIUGLTKNACELGAHGIRVNCISPPFGVATPMLINAWROGHDASTDADDAD	262
Qy		226	-----DKEEKSMKETLR-----IRRIGEPCDCAIGVSFLCSEDASYITGETV	268
Db		263	IDLDIAVPDSQOEKVKEEWVRGLATLKCATLR---PRDIAEAAFLLASDDSDRYISGNHL	318
Qy		269	VVGGG-TPSR	277
Db		319	VVDGGVTSTR	328

```

US-09-363-189B-6
; Sequence 6, Application US/09363189B
; Patent No. 6242228
; GENERAL INFORMATION:
; APPLICANT: SUGIYAMA, MASAKAZU
; APPLICANT: TONOUCHI, NAOTO
; APPLICANT: SUZUKI, SHUNICHI
; APPLICANT: YOKOZAKI, KENZO
; TITLE OF INVENTION: XYLIOTOL DEHYDRATION
; FILE REFERENCE: 0010-1024-0
; CURRENT APPLICATION NUMBER: US/09/363-189B-6
; CURRENT FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: JP10-210000
; PRIOR FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patent version 3.0

```

53, 189B
5047

```

; SEQ ID NO 6
; LENGTH: 262
; TYPE: PRT
; ORGANISM: Gluconobacter oxydans
US-09-363-189B-6

Query Match 23.2%; Score 324; DB 4; Length 262;
Best Local Similarity 30.1%; Pred. No. 1.8e-27;
Matches 78; Conservative 55; Mismatches 104; Indels 22; Gaps 3;

QY 33 KVALVTASTDGTGFATARRLAODGAHVHVVSSRRKQNVQAVATLOGEGLSVGTGVCHVK 92
DB 8 KVCVLTGAGNIGLATALALAEETGTAIALDMMNRALEKAEASVEKGEARSVCVDTVS 67
QY 93 AEDRRLVATAVKLHGIDILVSNAAVNPFFGSIIMDVTEEVMDTKLDINVKAPALMTKAV 152
DB 68 EEAIVGTVDVVRDCKIDFLFNAGYOGAFAPVDYFSDDFARVLTINVTGAFHLKAV 127
QY 153 VPEMERKGGSVVIVSSIAAFSPGFPYNNVSKTALLGLTKTLAIELAPRNIIVNCLAP 212
DB 128 SRQMTQNTGRIVNTASMAVGKPPNMAAYGASKGAIITETAALDLAPYNIIRVNAISP 187
QY 213 GLIKTSFRLMNDKEKE-----ESMKETLRIIRRLGEPEDCAGIVSF 254
DB 188 GYMGPGF---MM-EKQVELQAKVGOYFSTDPKVVAAQMGVPMRYGDIINEIPGVVAF 243
QY 255 LCSSEDASYITGETVVVGGG 273
DB 244 LLDGDSFMTGVNLPAGG 262

RESULT 7
US-08-241-766-13
; Sequence 13, Application US/08241766
; Patent No 5685590
; GENERAL INFORMATION:
; APPLICANT: JACOBS, W. R.
; APPLICANT: COLLINS, D. M.
; APPLICANT: BANERJEE, A.
; APPLICANT: GELISLE, G. W.
; APPLICANT: WILSON, T. M.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTING
; TITLE OF INVENTION: AND TREATING MYCOBACTERIAL INFECTIONS USING AN INHA AGENT
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FORSTER
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/241,766
; FILING DATE: 12-MAY-1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: MONROY, GLADYS H.
; REGISTRATION NUMBER: 32,430
; REFERENCE/DOCKET NUMBER: 25237-20003.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 484-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 247 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-241-766-13

Query Match 22.9%; Score 321; DB 1; Length 247;
Best Local Similarity 31.9%; Pred. No. 3.6e-27;
Matches 79; Conservative 53; Mismatches 100; Indels 16; Gaps 5;

QY 27 RDLANKVALVYASTDGTGFATARRLAODGAHVHVVSSRRKQNVQAVATLOGEGLSVTGT 86
DB 10 KPFEVSRSLVTVGGNRGIGLATAQLAADGHKVAVTHRG-----SGAPKGLFCV 58
QY 87 VCHVGKAEERLRVATAVKLH--GGIDILVSNAAVNPFFGSIIMDVTEEVMDTKLDINVKAP 145
DB 59 ECDVTDSDAVDR-AFTAVEEHQGPVEVLVSNAGLSA-DAFLMRMTEEFKFKVINANLTGA 116
QY 146 ALMTKAVVPEMERKGGSVVIVSSIAAFSPGFPYNNVSKTALLGLTKTLAIELAPRNI 205
DB 117 FRVAQRASRSRMORNKFGRMIFIGSVSGWIGNQANYAASKAGVIGMARSIARELSKANV 176
QY 206 RVNCLAPGLIKTSFRLMNDKEKEKMKETLRIIRRLGEPEDCAGIVSFCESEDASYITG 265
DB 177 TANVAPGGVIDTDMTRAL--DERIQOQALQFIPAKRVGTPAEVAGVSVFLASEDASYISG 234
QY 266 ETVVVG 273
DB 235 AVTPVDG 242

RESULT 8
US-08-375-962B-13
; Sequence 13, Application US/08375962B
; Patent No 5731195
; GENERAL INFORMATION:
; APPLICANT: SIMON, ANDRAS; HELLMAN, ULF; WERNSTEDT,
; APPLICANT: CRISTER; ERIKSSON, ULF.
; TITLE OF INVENTION: Isolated Nucleic Acid Molecule
; TITLE OF INVENTION: Which Codes for A 32 kDa Protein Having 11-CIS Retinol
; TITLE OF INVENTION: Dehydrogenase Activity, and Which Associates With P53
; TITLE OF INVENTION: a Portion of a Retinol Binding Protein Receptor
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect (ASCII standard)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/375,962B
; FILING DATE: 20-January-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/258,418
; FILING DATE: 6-October-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Pasqualini, Patricia A.
; REGISTRATION NUMBER: 34,894
; REFERENCE/DOCKET NUMBER: LUD 5372.1 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 244 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

```

```

; SEQ ID NO 6
; LENGTH: 262
; TYPE: PRT
; ORGANISM: Gluconobacter oxydans
US-09-363-189B-6

Query Match 23.2%; Score 324; DB 4; Length 262;
Best Local Similarity 30.1%; Pred. No. 1.8e-27;
Matches 78; Conservative 55; Mismatches 104; Indels 22; Gaps 3;

QY 33 KVALVTASTDGTGFATARRLAODGAHVHVVSSRRKQNVQAVATLOGEGLSVGTGVCHVK 92
DB 8 KVCVLTGAGNIGLATALALAEETGTAIALDMMNRALEKAEASVEKGEARSVCVDTVS 67
QY 93 AEDRRLVATAVKLHGIDILVSNAAVNPFFGSIIMDVTEEVMDTKLDINVKAPALMTKAV 152
DB 68 EEAIVGTVDVVRDCKIDFLFNAGYOGAFAPVDYFSDDFARVLTINVTGAFHLKAV 127
QY 153 VPEMERKGGSVVIVSSIAAFSPGFPYNNVSKTALLGLTKTLAIELAPRNIIVNCLAP 212
DB 128 SRQMTQNTGRIVNTASMAVGKPPNMAAYGASKGAIITETAALDLAPYNIIRVNAISP 187
QY 213 GLIKTSFRLMNDKEKE-----ESMKETLRIIRRLGEPEDCAGIVSF 254
DB 188 GYMGPGF---MM-EKQVELQAKVGOYFSTDPKVVAAQMGVPMRYGDIINEIPGVVAF 243
QY 255 LCSSEDASYITGETVVVGGG 273
DB 244 LLDGDSFMTGVNLPAGG 262

RESULT 7
US-08-241-766-13
; Sequence 13, Application US/08241766
; Patent No. 5686590
; GENERAL INFORMATION:
; APPLICANT: JACOBS, W. R.
; APPLICANT: COLLINS, D. M.
; APPLICANT: BANERJEE, A.
; APPLICANT: GELISLE, G. W.
; APPLICANT: WILSON, T. M.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR DETECTING
; TITLE OF INVENTION: AND TREATING MYCOBACTERIAL INFECTIONS USING AN INHA AGENT
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FORSTER
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/241,766
; FILING DATE: 12-MAY-1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: MONROY, GLADYS H.
; REGISTRATION NUMBER: 32,430
; REFERENCE/DOCKET NUMBER: 25237-20003.20
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 484-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 247 amino acids
; TYPE: amino acid
; TOPOLOGY: linear

; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-241-766-13

Query Match 22.9%; Score 321; DB 1; Length 247;
Best Local Similarity 31.9%; Pred. No. 3.6e-27;
Matches 79; Conservative 53; Mismatches 100; Indels 16; Gaps 5;

QY 27 RDLANKVALVYASTDGTGFATARRLAODGAHVHVVSSRRKQNVQAVATLOGEGLSVTGT 86
DB 10 KPFEVSRSLVTVGGNRGIGLATAQLAADGHKVAVTHRG-----SGAPKGLFCV 58
QY 87 VCHVGKAEERLRVATAVKLH--GGIDILVSNAAVNPFFGSIIMDVTEEVMDTKLDINVKAP 145
DB 59 ECDVTDSDAVDR-AFTAVEEHQGPVEVLVSNAGLSA-DAFLMRMTTEEFKVINANLTGA 116
QY 146 ALMTKAVVPEMERKGGSVVIVSSIAAFSPGFPYNNVSKTALLGLTKTLAIELAPRNI 205
DB 117 FRVAQRASRSRMORNKFGRMIFIGSVSGWIGNQANYAASKAGVIGMARSIARELSKANV 176
QY 206 RVNCLAPGLIKTSFRLMNDKEKEKMKETLRIIRRLGEPEDCAGIVSFCESEDASYITG 265
DB 177 TANVAPGVITDMDTRAL--DERIQOQALQFIPAKRVGTPAEVAGVSVFLASEDASYISG 234
QY 266 ETVVVG 273
DB 235 AVTPVDG 242

RESULT 8
US-08-375-962B-13
; Sequence 13, Application US/08375962B
; Patent No. 5731195
; GENERAL INFORMATION:
; APPLICANT: SIMON, ANDRAS; HELLMAN, ULF; WERNSTEDT,
; APPLICANT: CRISTER; ERIKSSON, ULF.
; TITLE OF INVENTION: Isolated Nucleic Acid Molecule
; TITLE OF INVENTION: Which Codes for A 32 kDa Protein Having 11-CIS Retino
; TITLE OF INVENTION: Dehydrogenase Activity, and Which Associates With P53
; TITLE OF INVENTION: a Portion of a Retinol Binding Protein Receptor
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage
; COMPUTER: IBM
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect (ASCII standard)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/375,962B
; FILING DATE: 20-January-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/258,418
; FILING DATE: 6-October-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Pasqualini, Patricia A.
; REGISTRATION NUMBER: 34,894
; REFERENCE/DOCKET NUMBER: LUD 5372.1 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 244 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

```

MOLECULE TYPE: protein

FEATURE:

NAME/KEY: E.coli 3-oxoacyl[acyl carrier protein]

NAME/KEY: reductase (FABG)

US-08-375-962B-13

Query Match 22.4% Score 313; DB 1; Length 244;
Best Local Similarity 32.8%; Pred. No. 2.6e-26;
Matches 81; Conservative 53; Mismatches 95; Indels 18; Gaps 6;

QY 33 KVALVTASTDIGFAIARRLAQDGAHVSVSRKQONVDQAVAT---LOGEGLSVGTGVCH 89

Db 6 KIALVTGASRGIGRAIAETLAARGCK-VIGTATSENGAQALSDYLGANGKGLMLNVT--- 61

QY 90 VGKAEDRELRLVATAVKLH---GGIDILVSNAAVNPFFGSGIMDVTEEVWDKTLIDINVKAPA 146

Db 62 -----DPASIESVLEKIRAEFGEVDILVNNAGITR-DNLLMRMKDEEWNDDIETNLSVVF 115

QY 147 LMTKAVVPMEKRGSGVSVVSSIAAFSPSPGFPYNNVSKTALLGLTKTLAIELAPRNIR 206

Db 116 RLSKAVRAMMKRHRGRIITIGSVGTMGNGGQANYAAKAGLIGFSKSLAREVASRGIT 175

QY 207 VNCIAPGLIKTSFRMLMMDKEESKMETLRIIRLGEPEDCAGIVSFLCSEDASYITGE 266

Db 176 VNVVAPGFETDMTRAL---SDDQRAGILAQVPAGRLGGAQEIANAVAFASDEAAAYITGE 233

QY 267 TVVVG 273

Db 234 TLHVNGG 240

RESULT 9

US-08-562-114B-13

Sequence 13, Application US/08562114B

Patent No. 5972646

GENERAL INFORMATION:

APPLICANT: ERIKSSON ET AL.

TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULE WHICH CODES FOR A

TITLE OF INVENTION: 32 KDA PROTEIN HAVING 11-CIS RETINOL HYDROGENASE

TITLE OF INVENTION: ACTIVITY, AND WHICH ASSOCIATES WITH P63, A

TITLE OF INVENTION: PORTION OF A RETINOL BINDING PROTEIN RECEPTOR

NUMBER OF SEQUENCES: 31

CORRESPONDENCE ADDRESS:

ADDRESSEE: Felfe & Lynch

STREET: 805 Third Avenue

CITY: New York City

STATE: New York

ZIP: 10022

COMPUTER READABLE FORM: 3.5 inch, 144 kb storage

MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage

COMPUTER: IBM

OPERATING SYSTEM: PC-DOS

SOFTWARE: Wordperfect 5.1 and ASCII

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/562,114B

FILING DATE: 22-No. 5972646ember-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/375,962

FILING DATE: 20-January-1995

ATTORNEY/AGENT INFORMATION:

NAME: Kohli, Vineet

REGISTRATION NUMBER: 37,003

REFERENCE/DOCKET NUMBER: LUD 5372.2 CIP

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 688-9200

TELEFAX: (212) 688-3884

INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:

LENGTH: 244 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

FEATURE:

NAME/KEY: E.coli 3-oxoacyl[acyl carrier protein]reductase

NAME/KEY:

US-08-562-114B-13

Query Match 22.4% Score 313; DB 2; Length 244;
Best Local Similarity 32.8%; Pred. No. 2.6e-26;
Matches 81; Conservative 53; Mismatches 95; Indels 18; Gaps 6;

QY 33 KVALVTASTDIGFAIARRLAQDGAHVSVSRKQONVDQAVAT---LOGEGLSVGTGVCH 89

Db 6 KIALVTGASRGIGRAIAETLAARGCK-VIGTATSENGAQALSDYLGANGKGLMLNVT--- 61

QY 90 VGKAEDRELRLVATAVKLH---GGIDILVSNAAVNPFFGSGIMDVTEEVWDKTLIDINVKAPA 146

Db 62 -----DPASIESVLEKIRAEFGEVDILVNNAGITR-DNLLMRMKDEEWNDDIETNLSVVF 115

QY 147 LMTKAVVPMEKRGSGVSVVSSIAAFSPSPGFPYNNVSKTALLGLTKTLAIELAPRNIR 206

Db 116 RLSKAVRAMMKRHRGRIITIGSVGTMGNGGQANYAAKAGLIGFSKSLAREVASRGIT 175

QY 207 VNCIAPGLIKTSFRMLMMDKEESKMETLRIIRLGEPEDCAGIVSFLCSEDASYITGE 266

Db 176 VNVVAPGFETDMTRAL---SDDQRAGILAQVPAGRLGGAQEIANAVAFASDEAAAYITGE 233

QY 267 TVVVG 273

Db 234 TLHVNGG 240

RESULT 10

US-08-729-594A-13

Sequence 13, Application US/08729594A

Patent No. 6280997

GENERAL INFORMATION:

APPLICANT: Eriksson, Ulf; Simon, Andras; Romert, Anna

TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULE WHICH

TITLE OF INVENTION: CODES FOR A 32 KDA PROTEIN HAVING 11-CIS RETINOL DEHYDROGEN

TITLE OF INVENTION: ACTIVITY, AND WHICH ASSOCIATES WITH P63, A PORTION OF A

TITLE OF INVENTION: RETINOL BINDING PROTEIN RECEPTOR

NUMBER OF SEQUENCES: 41

CORRESPONDENCE ADDRESS:

ADDRESSEE: Felfe & Lynch

STREET: 805 Third Avenue

CITY: New York City

STATE: New York

ZIP: 10022

COMPUTER READABLE FORM: 3.5 inch, 144 kb storage

MEDIUM TYPE: Diskette, 3.5 inch, 144 kb storage

COMPUTER: IBM

OPERATING SYSTEM: PC-DOS

SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/729,594A

FILING DATE: 11-October-1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/562,114

FILING DATE: 22-No. 6280997ember-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/375,962

FILING DATE: 20-January-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/258,418

FILING DATE: 10-June-1994

ATTORNEY/AGENT INFORMATION:

NAME: Hanson, No. 6280997man D.

REGISTRATION NUMBER: 30,946

REFERENCE/DOCKET NUMBER: LUD 5372.3

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 688-9200

TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 244 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: E.coli 3-oxoacyl[acyl carrier protein]reductase
NAME/KEY:
US-08-729-594A-13

Query Match 22.48; Score 313; DB 4; Length 244;
Best Local Similarity 32.88; Pred. No. 2.6e-26;
Matches 81; Conservative 53; Mismatches 95; Indels 18; Gaps 6;
QY 33 KVALVTASTDGGIGFAIRRLAODGAHVVSRRKQONVDQAVAT---LOGEGSLVTGTVCH 89
DB 6 KIALVTGASRGIGRAIAETLAAGGK-VIGTATSENGAQALSYLGANGKGLMLNVT--- 61
QY 90 VGKAEDRELRVATAVKLH---GGIDILVSNAAVNPFFGSDMTVEEVMDKTLIDINVKAPA 146
DB 62 -----DPASIESVLEKIRAECEVDILVNNAGITR-DNLMRMKDEEWNDIETNLSVF 115
QY 147 LMTKAVVPEMEKRGSGSVVIVSSIAAFSPSPGFPYNNVSKTALLGLTKTLAIELAPRNR 206
DB 116 RLSKAVVRAMMKRHRITIGSVVTMGNGGQANYAAKAGLIGFSKSLAREVASRGIT 175
QY 207 VNCIAPGLIKTSFMRMLMDKEESMKETLIRRLGEPEDCAGIVFLCSEDASYITGE 266
DB 176 VNVVAPGFIEDMTAL--SDQKAGILAQVPAQRLGGAQEIANAVAFSLASDEAAAYITGE 233
QY 267 TVVVGSG 273
DB 234 TLVNGG 240

RESULT 11
US-08-594-808B-7
Sequence 7, Application US/08594808B
Patent No. 5804423
GENERAL INFORMATION:
APPLICANT: Klasen, Ralf
APPLICANT: Bringer-Meyer, Stephanie
APPLICANT: Sahm, Hernand
APPLICANT: Hollenberg, Cornelia P
TITLE OF INVENTION: MICROBIOLOGICAL METHOD OF MAKING
TITLE OF INVENTION: 5-KETOGLUCONATE
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: The Firm of Karl F. Ross, PC
STREET: 5676 Riverdale Ave.
CITY: Bronx
STATE: New York
COUNTRY: USA
ZIP: 10471-0900
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/594,808B
FILING DATE: 07-FEB-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Myers, Jonathan E
REGISTRATION NUMBER: 26,963
REFERENCE/DOCKET NUMBER: 19893
TELECOMMUNICATION INFORMATION:
TELEPHONE: (718) 884-6600
TELEFAX: 718/601-1099

TELEX: 620428
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 256 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-594-808B-7
Query Match 22.38; Score 312.5; DB 1; Length 256;
Best Local Similarity 31.28; Pred. No. 3.2e-26;
Matches 77; Conservative 53; Mismatches 110; Indels 7; Gaps 3;
QY 30 LANKVALVTASTDGGIGFAIRRLAODGAHVVSRRKQONVDQAVATLOGEGSLVTGTVCH 89
DB 9 LSGARALVTGASRGIGITLAKGLARYGAEEVVLNGRNAESLDSAQSGFEAGKASTAVFD 68
QY 90 VGKAEDRELR---VATAVKLHGGIDILVSNAAVNPFFGSDMTVEEVMDKTLIDINVKAPA 146
DB 69 V---TDQDAVIDGVAATERDMPIDILINNAGIQR-RAPLEEFKRDWDDLMSTNNVAVF 124
QY 147 LMTKAVVPEMEKRGSGSVVIVSSIAAFSPSPGFPYNNVSKTALLGLTKTLAIELAPRNR 206
DB 125 FVGQAVARHMIPIRGKIVNICSQSELARPGIAPYATKGAVKNLTKGMATDWRHGLQ 184
QY 207 VNCIAPGLIKTSFMRMLMDKEESMKETLIRRLGEPEDCAGIVFLCSEDASYITGE 266
DB 185 INGLAPGYFATEMTERLVADEEFTDNLCKRTPAGRWQGVQVELVGAAVFLSSRASFEVNGQ 244
QY 267 TVVVGSG 273
DB 245 VLMVDGG 251

RESULT 12
US-08-440-856A-3
Sequence 3, Application US/08440856A
Patent No. 5750873
GENERAL INFORMATION:
APPLICANT: DELLAPORTA, STEPHEN L.
TITLE OF INVENTION: MATERIALS AND METHODS FOR PRODUCING
TITLE OF INVENTION: PLANTS WITH SINGLE-SEX FLOWERS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 PENNSYLVANIA AVE. N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/440,856A
FILING DATE: 15-MAY-1995
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: MILLMAN, ROBERT A.
REGISTRATION NUMBER: 36,217
REFERENCE/DOCKET NUMBER: 05463-20001.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1517
TELEFAX: (202) 887-0763
TELEX: 706141
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 337 amino acids
TYPE: amino acid
STRANDEDNESS: single

QY 148 MTKAVPEMEKRGSGVIVSSIAAFSPGSPYNVSKTALLGLTKTLAIELAPRNIRV 207
Db 118 TPAALPGNRKAGRWIRIVASAHGLTASPYKSAVAAKHGVGFTKVTALTAGKGITC 177
QY 208 NCLAPGLIKTSF-----SRMLMDKEESMKETL-----RIRRLGEPEDCAGIVSFCLCS 257
Db 178 NAICPGYVLTPLVEAQIPDKMAHMDRETIVREYMLDRQPSRQFATTGQIGGTVVFLCS 237
QY 258 EDASVITGETVVVGGG 273
Db 238 GAADQITGTTISVDGG 253

RESULT 15
US-08-762-129-1
; Sequence 1, Application US/08762129
; Patent No. 5756299
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: A NOVEL HUMAN CARBONYL REDUCTASE
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08762,129
; FILING DATE: Herewith
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0171 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 244 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: Consensus
; CLONE: Consensus
US-08-762-129-1

Query Match 21.1%; Score 295; DB 1; Length 244;
Best Local Similarity 32.7%; Pred. No. 2.5e-24;
Matches 81; Conservative 44; Mismatches 107; Indels 16; Gaps 5;
QY 30 LANKVALVTASTDGIQFAIARRLAQDGAHVYVSSRKQONVDQAVATLQGGSLSVGTGVCH 89
Db 5 LAGRRVLVTGAGKIGRGTVQALHATGARVAVVSRQTQADLDSLVRCPG----IEPVCDV 60
QY 90 VGRAEDRERLVATKVLHGIDILVNAAV----NPFEGSIMDYETEVEWVKTLINVKAPA 146
Db 61 LGDWEATERALGSV----GPVDLLVNNAAVALLQPF----LEVTKEAFDRSFEVNLRAVI 112

QY 147 LMTKAVPEMEKRG--GGSVIVSSIAAFSPGSPYNVSKTALLGLTKTLAIELAPRNIRV 205
Db 113 QVSOIVARGLLIARGVPGAIVNVSSQCSQORAVTNHVSVCSTKGALDMLTKVMALELGPHKI 172
QY 206 RVNCLAPGLIKTSFSPSRMLMDKEESMKETLIRIRRLGEPEDCAGIVSFCLCEDASYITG 265
Db 173 RVNAVNPVTVMVTSMGOATWSDPHKAKTMXNRIPXGKFAEVEHVVNAILFLLSDRSGMTTG 232
QY 266 ETVVVGGG 273
Db 233 STLPVEGG 240

Search completed: January 22, 2002, 15:19:24
Job time: 114 sec

Result No.	Score	Query Match	Length	DB ID	Description
1	824.5	58.9	257	2	S66665
2	636	45.5	257	2	T32002
3	614	43.9	260	2	A2676
4	588.5	42.1	254	2	E95069
5	459.5	32.8	260	2	E70881
6	409	29.2	255	2	D83416
7	399	28.5	246	2	A59621
8	396.5	28.3	246	2	H72219
9	388	27.7	246	2	C83961
10	385	27.5	248	2	H70447
11	381.5	27.3	282	2	C71204
12	381	27.2	255	2	H84288
13	374	26.3	243	2	F86721
14	373.5	26.7	251	2	F23389
15	371	26.5	263	2	C75217
16	370	26.4	249	2	B83767
17	369.5	26.4	285	2	D85885
18	368.5	26.3	285	1	A65017
19	368.5	26.3	320	2	S22450
20	366	26.2	261	2	S35196
21	365.5	26.1	248	2	F69868
22	364.5	26.1	260	2	B48674
23	364	26.0	254	2	G75333
24	360.5	25.8	261	2	S54815
25	360	25.7	272	1	S05397
26	358.5	25.6	256	2	F70637
27	357	25.5	248	2	F81971
28	357	25.5	248	2	E81026
29	357	25.5	273	1	A48674
30	357	25.5	273	1	A48674

Db 193 LGLTRTLELAPKDIRVNCVVGIIKTDFSKVFHGNESLWKNFKHHQLQRIGESDCA 252
Qy 250 GIVSFLCSDASYITGVTVGG 272
Db 253 GIVSFLCSPDASYVNGENIAVAG 275
RESULT 2
T32002
hypothetical protein F36H9.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 31-Jan-2000
C:Accession: T32002
R:Dante, M.; Kramer, J.; Twyman, B.
submitted to the EMBL Data Library, July 1997
A:Description: The sequence of C. elegans cosmid F36H9.
A:Reference number: T22110
A:Accession: T32002
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-257 <DNA>
A:Cross-references: EMBL:AF016668; PIDN:AAB66092.1; GSPDB:GN00020; CESP:F36H9.3
A:Experimental source: strain Bristol N2; clone F36H9
C:Genetics:
A:Gene: CESP:F36H9.3
A:Map position: 2
A:Introns: 21/3; 49/3; 192/1
C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match 45.5%; Score 636; DB 2; Length 257;
Best Local Similarity 50.6%; Pred. No. 2.8e-4;
Matches 126; Conservative 49; Mismatches 74; Indels 0; Gaps 0;
Qy 30 LANKVALVTASTDGGIGFAIARRLAQDGAHVSVSRKQNVDOAVATLOGEGLSVTGVCH 89
Db 9 LTRVALVTASTKGIGFAIARQLGARGASVVCSSKKNVDEAVAALENDIAHGHTAH-68
Qy 90 VGKADRELRVATAVKLHGGDILVSNAAVNPFFGSIIMDVTEVWDKPLDINVKAPALMT 149
Db 69 VGKSDRTKLDFTRDKLIDLVSNAAVNPFFGSIIMDVTEVWDKPLDINVKAPALMT 128
Qy 150 KAVVPEMEKRGGSVVIVSSIAAFSPSPFPYNSKTAALLGLTKTALAIELAPNRVNC 209
Db 129 KEAVPHLEASGRNVFVSSVAGSPMNEIGAYSMKTTLLGLSKSCALNLARNRNVNS 188
Qy 210 LAPGLIKTSFRMLWMDKEESKMETLRIRRLGEPEDCAGIVSFLCSDASYITGVTV 269
Db 189 IAPGIITDFQSVLFSEDESKQWLSQIAQRFRFGDPDECAVAFVLSDEASYISGETIG 248
Qy 270 VGGTTPSRL 278
Db 249 INGMHARI 257
RESULT 3
T22676
hypothetical protein F54F3.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 31-Jan-2000
C:Accession: T22676; T24959
R:Perocy, C.; Lloyd, C.
submitted to the EMBL Data Library, September 1996
A:Reference number: Z19598
A:Accession: T22676
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-260 <ML>
A:Cross-references: EMBL:Z79696; PIDN:CAB01974.1; GSPDB:GN00023; CESP:F54F3.4
A:Experimental source: clone F54F3
R:Lloyd, C.
submitted to the EMBL Data Library, November 1996
A:Reference number: Z19960

Db 193 LGLTRTLELAPKDIRVNCVVGIIKTDFSKVFHGNESLWKNFKHHQLQRIGESDCA 252
Qy 250 GIVSFLCSDASYITGVTVGG 272
Db 253 GIVSFLCSPDASYVNGENIAVAG 275
RESULT 2
T32002
hypothetical protein F36H9.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 31-Jan-2000
C:Accession: T32002
R:Dante, M.; Kramer, J.; Twyman, B.
submitted to the EMBL Data Library, July 1997
A:Description: The sequence of C. elegans cosmid F36H9.
A:Reference number: T22110
A:Accession: T32002
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-257 <DNA>
A:Cross-references: EMBL:AF016668; PIDN:AAB66092.1; GSPDB:GN00020; CESP:F36H9.3
A:Experimental source: strain Bristol N2; clone F36H9
C:Genetics:
A:Gene: CESP:F36H9.3
A:Map position: 2
A:Introns: 21/3; 49/3; 192/1
C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match 45.5%; Score 636; DB 2; Length 257;
Best Local Similarity 50.6%; Pred. No. 2.8e-4;
Matches 126; Conservative 49; Mismatches 74; Indels 0; Gaps 0;
Qy 30 LANKVALVTASTDGGIGFAIARRLAQDGAHVSVSRKQNVDOAVATLOGEGLSVTGVCH 89
Db 9 LTRVALVTASTKGIGFAIARQLGARGASVVCSSKKNVDEAVAALENDIAHGHTAH-68
Qy 90 VGKADRELRVATAVKLHGGDILVSNAAVNPFFGSIIMDVTEVWDKPLDINVKAPALMT 149
Db 69 VGKSDRTKLDFTRDKLIDLVSNAAVNPFFGSIIMDVTEVWDKPLDINVKAPALMT 128
Qy 150 KAVVPEMEKRGGSVVIVSSIAAFSPSPFPYNSKTAALLGLTKTALAIELAPNRVNC 209
Db 129 KEAVPHLEASGRNVFVSSVAGSPMNEIGAYSMKTTLLGLSKSCALNLARNRNVNS 188
Qy 210 LAPGLIKTSFRMLWMDKEESKMETLRIRRLGEPEDCAGIVSFLCSDASYITGVTV 269
Db 189 IAPGIITDFQSVLFSEDESKQWLSQIAQRFRFGDPDECAVAFVLSDEASYISGETIG 248
Qy 270 VGGTTPSRL 278
Db 249 INGMHARI 257
RESULT 3
T22676
hypothetical protein F54F3.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 31-Jan-2000
C:Accession: T22676; T24959
R:Perocy, C.; Lloyd, C.
submitted to the EMBL Data Library, September 1996
A:Reference number: Z19598
A:Accession: T22676
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-260 <ML>
A:Cross-references: EMBL:Z79696; PIDN:CAB01974.1; GSPDB:GN00023; CESP:F54F3.4
A:Experimental source: clone F54F3
R:Lloyd, C.
submitted to the EMBL Data Library, November 1996
A:Reference number: Z19960

A:Accession: T24959
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-260 <W12>
A:Cross-references: EMBL:Z81592; PIDN:CAB04734.1; GSPDB:GN00023; CESP:F54F3.4
A:Experimental source: clone T16G1
C:Genetics:
A:Gene: CESP:F54F3.4
A:Map position: 5
A:Introns: 21/2; 81/3; 201/3
C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match 43.9%; Score 614; DB 2; Length 260;
Best Local Similarity 48.0%; Pred. No. 1.6e-41;
Matches 120; Conservative 56; Mismatches 70; Indels 4; Gaps 2;
Qy 33 KVALVTASTDGGIGFAIARRLAQDGAHVSVSRKQNVDOAVATLOGEGLS-VTGTVCVHG 91
Db 11 KVAIVTAATKGIGLAIAERLDEGASVVIGSRNQKNVDEAIEYLNKGLTKVAGIAGHA 70
Qy 92 KAEDRELRVATAVKLHGGDILVSNAAVNPFFGSIIMDVTEVWDKPLDINVKAPALMTKA 151
Db 71 STDQKKLVDTFLOKFGKINILVNNHGINPAFGHILEVSDQVNDKLFVNVKAGFQMTKL 130
Qy 152 VVPEMEKRGGSVVIVSSIAAFSPSPFPYNSKTAALLGLTKTALAIELAPNRVNCIA 211
Db 131 VPHIAKEGGGAIIFNASYSAYKSPPGIAAYGVTKTLVGLTRALAMGLAKDNIRVNGIA 190
Qy 212 PGLIKTSFRMLW---MDKEESKMETLRIRRLGEPEDCAGIVSFLCSDASYITGVTV 268
Db 191 PCVTKTKMSQVLPDGGDAEKELTDIOEIALGRLVDPDCACTVAYLASDDSSYITGEMI 250
Qy 269 VVGGTTPSRL 278
Db 251 ILAGVQARL 260
RESULT 4
E85069
hypothetical protein AT4g05530 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001
C:Accession: E85069
R:anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Sp
Nature 402, 769-777, 1999
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A:Reference number: A85001; MUID:20083488
A:Accession: E85069
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-254 <STO>
A:Cross-references: GB:NC_001268; NID:g7267313; PIDN:CAB81095.1; GSPDB:GN00140
C:Genetics:
A:Gene: AT4g05530
A:Map position: 4

Query Match 42.1%; Score 588.5; DB 2; Length 254;
Best Local Similarity 50.6%; Pred. No. 1.7e-39;
Matches 126; Conservative 35; Mismatches 85; Indels 3; Gaps 2;
Qy 30 LANKVALVTASTDGGIGFAIARRLAQDGAHVSVSRKQNVDOAVATLOGEGLSVTGVTVCH 89
Db 9 LEGKVAIVTASTQIGFGITERFGLGASVVVSRKQNVDEAVAKLKSCKIDAYGIYCH 68
Qy 90 VGKADRELRVATAVKLHGGDILVSNAAVNPFFGSIIMDVTEVWDKPLDINVKAPALMT 149
Db 69 VSNQAHRLNVEKTVQYKGIKIDIVVCNAAANPSTDPILSSKEAVLDKLNWEINVSSILL 128
Qy 150 KAVVPEMEKRGGSVVIVSSIAAFSPSPFPYNSKTAALLGLTKTALAIELAPNRVNC 209
Db 129 QDMAPHEK--GSSVIFITSIAGSPGAMAMYGVTALLGLTKALAAEMAP-DTRVNA 185

A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-255 <STO>
A;Cross-references: GB:AE004609; GB:AE004031; NID:g9947810; PIDN:AAG05217.1; GSPDB:CN
A;Experimental source: Strain PA01
C;Genetics:
A;Gene: PA1828
C;Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match	29.2%;	Score	409;	DB	2;	Length	255;
Best Local Similarity	37.3%;	Pred.	No. 3e-25;				
Matches	91;	Conservative	45;	Mismatches	108;	Indels	0; Gaps 0;
QY	30	LANKVALVTASTDGGIFATARRLAODGAHVHVSRRKQNVDOQAVATLOGEGLSVTTGVCH	89				
		: :	:	:	:	:	:
Db	9	LDGKIAPVSGASRGIGETAKLLAQOAGAHIVSSRKIDGCQAVADAITAEGGKATAIACH	68				
		:	:	:	:	:	:
QY	90	VGRAEDRERLIVATAVLHGSGIDLTVSNAAVNPPFGSIMDVTEEVMDTKTDLIDINVKAPALWT	149				
		:	:	:	:	:	:
Db	69	IGEMEQIQNVFAIRQFOGRDLTLVNAATNPQFCNWLETDLGAFOKFTVDVNRIGYYFMS	128				
		:	:	:	:	:	:
QY	150	KAYVPWEKRGGSVVIVSIAAFSPGSPFNYSKTAALLGKTLTALIELAPRNIRVNC	209				
		:	:	:	:	:	:
Db	129	IECGKLMEKHGGGSIINVASINGVSGEQQIYSVTKAAVISMTKYFAECAQFGTRCNA	188				
		:	:	:	:	:	:
QY	210	LAPGLIKTSFRMLMWDKEEKESMKTLRIIRRLGEDPCAGIVSFPLCSEDASYITGETVW	269				
		:	:	:	:	:	:
Db	189	LLPGLTDTNKFASALVKNDAIRNLALQRILPKRYAEPSSEMAGAVLYLASEASSYTTGVALN	248				
		:	:	:	:	:	:
QY	270	VGGG	273				
		:	:				
Db	249	VDGG	252				

RESULT 7
 A369621 3-oxoacyl-[acyl-carrier-protein] reductase (EC 1.1.1.100) [validated] - Bacillus subtilis
 N:Alternate names: 3-ketoacyl-acyl carrier protein reductase
 C:Species: Bacillus subtilis
 C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 21-Jul-2000
 C:Accession: A369621; PC01176; T46633
 C:Authors: R. Ogasawara, N. Moszer, I. Albertini, A.M. Alloni, G. Azevedo, V. Ber
 C.: Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber
 C.: Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;
 A.: Ehrlich, S.D.; Emmerison, P.T.; Entian, K.D.; Errington, J.; Fabre, C.; Ferrari,
 Nature 390, 249-256, 1997
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gal
 iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.
 Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardino
 A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mau
 y, M.; Ogawa, K.; Ogihara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portete
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanl
 A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Se
 akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiya
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida
 A:Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
 A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis
 A:Reference number: A65680.
 A:Reference number: A65680. PMID: 98044033

A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-246 <GUN>
A;Cross-references: GB:26633902; PIDN:CABI3464.1; PID:g266339
A;Experimental source: strain 168
R;Oguro, A.; Kakeshita, H.; Takamatsu, H.; Nakamura, K.; Yamane, K.
Gene 172, 17-24, 1996
A;Title: The effect of Srb, a homologue of the mammalian SRP receptor alpha-subunit,
A;Reference number: JCA4819; MUID:96257247
A;Accession: PC4176
A;Molecule type: DNA
A;Residues: 230-246 <OGU>
A;Cross-references: DDBJ:D64116; NID:g1389548; PIDN:BAA10974.1; PID:g1237012
R;Cronan, J.E.; Morbidoni, H.R.; de Mendoza, D.

J. Bacteriol. 178, 4794-4800, 1996
 A:Title: Bacillus subtilis acyl carrier protein is encoded in a cluster of lipid biosyn-
 A:Reference number: 22107; MUID:96326321
 A:Accession: T46633
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-22, 'A', 24-246 <CRO>
 A:Cross-references: EMBL:059433; NID:gl502418; PIDN:AAC44307.1; PID:gl502421
 A:Experimental source: strain 168
 C:Genetics:
 A:Gene: fabG; srb
 A:Map position: 135-145 degrees
 C:Function:
 A:Description: EC 1.1.1.100 [validated; MUID:96326321]
 A:Pathway: fatty acid biosynthesis
 C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology
 C:Keywords: fatty acid biosynthesis; NADP; oxidoreductase
 F:5-185/Domain: short-chain alcohol dehydrogenase homology <SADH>
 F:154/Active site: Tyr #status predicted

Query Match 28.5%; Score 399; DB 2; Length 246;
 Best Local Similarity 37.6%; Pred. NO. 1.8e-24;
 Matches 92; Conservative 50; Mismatches 99; Indels 4; Gaps 3;

QY 30 LANKVALVTASTDGGIGFATARRLAQDGAHVYVS-SRQONVDQAVATLOGGLSVTGTVC 88
 DB 2 LNDKTAIVTGAASGIGKATLLFAQEGATVIAGDISKENLDSLVKEAGLPKGVDPVVLN 62
 QY 89 HGKAEDRERLAVATVAKLHGIDILVSNAAVNPFFGSIMDVTEVVDKTLIDINVKAPALM 148
 DB 62 DVSNPDEVQNMKETSLSVSTIDILVNNAGITR-DNLIMRKEDWDVNNINLKGVFNC 120
 QY 149 TRAVPEMEKRGGSVVIVSSIAAFSPGFPSPYNSKTLALLGTLKTALAIELAPRNIRVN 208
 DB 121 TRAVTRQMMKQSRGIINNVSVGSGNPGQANYAAKAGVIGLTKSSAKELASRNITVN 180
 QY 209 CLAPGLIKTSFMRMLWMDKEESMKETLIRRLGEPEDCAGIVSFLCSEDASYITGETV 268
 DB 181 ATAPGISIDMTDKL--AKVDQDEMLKQIPLARFGEPDSVSVVTFPLASEGARYMTGTL 238
 QY 269 VVGGG 273
 DB 239 HIDGG 243

RESULT 8
 H72219
 3-oxoacyl-(acyl carrier protein) reductase - Thermotoga maritima (strain MSB8)
 C:Species: Thermotoga maritima
 C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
 C:Accession: H72219
 R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey
 Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
 C.M.
 Nature 399, 323-329, 1999
 A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
 A:Reference number: A72200; MUID:99287316
 A:Accession: H72219
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-246 <ARN>
 A:Cross-references: GB:AE001811; GB:AE000512; NID:g4982291; PIDN:AAD36790.1; PID:g498230
 A:Experimental source: strain MSB8
 C:Genetics:
 A:Gene: TM1724
 C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology
 F:6-185/Domain: short-chain alcohol dehydrogenase homology <SAD>

Query Match 28.3%; Score 396.5; DB 2; Length 246;
 Best Local Similarity 36.5%; Pred. NO. 2.8e-24;
 Matches 89; Conservative 54; Mismatches 198; Indels 3; Gaps 3;

QY 30 LANKVALVTASTDGGIGFATARRLAQDGAHVYVS-SRQONVDQAVATLOGGLSVTGTVC 89
 DB 3 LEGKVLITGAASGIGKATLLFAQEGATVIAGDISKENLDSLVKEAGLPKGVDPVVLN 62
 QY 90 VGKAEDRERLAVATVAKLHGIDILVSNAAVNPFFGSIMDVTEVVDKTLIDINVKAPALMT 149
 DB 63 VTRDOIKEVKEVKVQKGRIDVLNNAGITR-DALLVRKMEEDMDAVINNLKGVNFVT 121
 QY 150 KAVVPEMEKRGGSVVIVSSIAAFSPGFPSPYNSKTLALLGTLKTALAIELAPRNIRVNC 209
 DB 122 QMVVPYMIKORNGSIYVSVVGIYNGPQTNAAKAGVIGMFKTWAKELAGNIRVNA 181
 QY 210 LAPGLIKTSFMRMLWMDKEESMKETLIRRLGEPEDCAGIVSFLCSEDASYITGETV 269
 DB 182 VAPGFIETPMTKEL-PEKARETALS-R-IPLGRFGKPEEVAQVILFLASDESSVVTGQVIG 239
 QY 270 VGGG 273
 DB 240 IDGG 243

RESULT 9
 C83961
 3-oxoacyl-(acyl-carrier protein) reductase fabG [imported] - Bacillus halodurans (str
 C:Species: Bacillus halodurans
 C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 01-Dec-2000
 C:Accession: C83961
 R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; H
 Nucleic Acids Res. 28, 4317-4331, 2000
 A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a
 A:Reference number: AB3650; MUID:20263314
 A:Accession: C83961
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-246 <STO>
 A:Cross-references: GB:AP001515; GB:BA000004; NID:g10174886; PIDN:BA806210.1; GSPDB:G
 A:Experimental source: strain C-125
 C:Genetics:
 A:Gene: fabG

Query Match 27.7%; Score 388; DB 2; Length 246;
 Best Local Similarity 37.9%; Pred. NO. 1.3e-23;
 Matches 94; Conservative 48; Mismatches 96; Indels 10; Gaps 4;

QY 30 LANKVALVTASTDGGIGFATARRLAQDGAHVYVS-SRQONVDQAVATLOGGLSVTGTVC 88
 DB 2 LOGKTAIVTGAASGIGKATLLFAQEGATVIAGDISKENLDSLVKEAGLPKGVDPVVLN 61
 QY 89 HGKAEDRERLAVATVAKLHGIDILVSNAAVNPFFGSIMDVTEVVDKTLIDINVKAP 145
 DB 62 DVADSESVQAMKEIDITDFGAVDILVNNAGITRDNLF----MRKEDMDAVIDNLKGV 117
 QY 146 ALMTKAVPEMEKRGGSVVIVSSIAAFSPGFPSPYNSKTLALLGTLKTALAIELAPRN 205
 DB 118 FHCRAVTRPMKQRFRIINNVSVVGVVAGNAGQANYAAKAGVIGLTKTLARELANRNI 177
 QY 206 RYNCIAPGLIKTSFMRMLWMDKEESMKETLIRRLGEPEDCAGIVSFLCSEDASYITG 265
 DB 178 TVNAVAPGFIETDGTGEL--PEDVKAQMLGQIPLARLQGPPEEVAKAVRFLASDDASYLTG 235
 QY 266 EVVVG 273
 DB 236 QTIHNGG 243

RESULT 10
 H70447
 3-oxoacyl-(acyl-carrier-protein) reductase (EC 1.1.1.100) - Aquifex aeolicus
 C:Species: Aquifex aeolicus
 C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 11-Jun-1999
 C:Accession: H70447

Qy	90	VGKAEDRLRVATAVKLLHGIGDILVSNAAVPPFGSIMDYTEEVWDKTTLDINVKAPALMT	149
Db	64	--AKDAQIVKKTVEFGRGLDILVNNAGIVP-YGNIETSEEDFDKTMVAVNVKGFLLS	119
Qy	150	KAVVPENKKRGGSVVILVSSIAFSPSPGSPYVSVKTKALLGUTKTALAPLNRIIVNC	209
Db	120	KYAYEQMKHKGGSYVNVSSGALIGIPRRCVYSVSKAALLGLTSLAVDYVDYGRVNA	179
Qy	210	LAPGLIKTS--FSRLMWMDKEESMKETLRI--RLGPEDCAGIVSFILCSEDASYITG	265
Db	180	VCPTTQSEGLMARVKASPNPELLKKMTSRIPMKRLGKEEETAFILFAACDEAGFWMT	239
Qy	266	ETVVVGGGT 274	
Db	240	SIINIDGGS 248	

RESULT 15
C75217
probable short-chain dehydrogenase/reductase PAB2177 - *Pyrococcus abyssi* (strain Orsa
C; Species: *Pyrococcus abyssi*
C; Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
C; Accession: C75217
R; anonymous, Genoscope
submitted to the EMBL Data Library, July 1999
A; Description: *Pyrococcus abyssi* genome sequence: insights into archaeal chromosome s
A; Reference number: A75001
A; Accession: C75217
A; Status: preliminary
A; Molecule type: DNA
A; Residues: 1-263 <RAW>
A; Cross-references: GB:AJ248283; GB:AL096836; NID:g5457433; PIDN:CAB49186.1; PID:g545
A; Experimental source: strain Orsay
C; Genetics:
A; Gene: PAB2177
C; Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology
F; 9-188/Domain: short-chain alcohol dehydrogenase homology <SAB>

Query Match 26.5%; Score 371; DB 2; Length 263;
Best Local Similarity 36.2%; Pred. NO. 3.2e-22;
Matches 92; Conservative 53; Mismatches 97; Indels 12; Gaps 5;

Search completed: January 22, 2002, 15:19:47
Job time: 102 sec

GenCore version: 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2002, 15:19:25; Search time 9.93 Seconds
(without alignments)
1026.467 Million cell updates/sec

Title: US-09-866-034-2

Perfect score: 1399

Sequence: 1 MHKAGLLGLCARAWNSVRMA.....DASYITGETVVVGCTPSRL 278

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 100059 segs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	800.5	57.2	257	1	HE27_HUMAN
2	399	28.5	246	1	FABG_BACSU
3	396.5	28.3	246	1	FABG_THEMA
4	385	27.5	248	1	FABG_AQUAE
5	379.5	27.1	260	1	TRN2_HYONI
6	373.5	26.7	251	1	Y325_THEMA
7	368.5	26.3	320	1	FABG_CUPLA
8	364.5	26.1	260	1	TRN2_DATST
9	362.5	25.9	263	1	UCPA_ECOLI
10	360	25.7	272	1	DHK1_STRVN
11	357	25.5	273	1	TRN1_DATST
12	354	25.3	241	1	FABG_RICPR
13	354	25.3	278	1	Y4LA_RHISN
14	349	24.9	256	1	Y019_THEMA
15	349	24.9	258	1	DHG2_BACSU
16	345	24.7	246	1	NODG_AZOB
17	343	24.5	244	1	FABG_VIBCH
18	342	24.4	247	1	FABG_SYNY3
19	341	24.4	248	1	FABG_CHLMO
20	340	24.3	242	1	FABG_HAEIN
21	337	24.1	247	1	FABG_CHLTR
22	336.5	24.1	256	1	BUDC_KLEPN
23	336.5	24.1	261	1	DHG2_BACME
24	336	24.0	245	1	NODG_RHIS3
25	335	23.9	261	1	DHKR_STRCM
26	334	23.9	267	1	HDHA_CLOSO
27	333.5	23.8	248	1	FABG_CHLPN
28	333.5	23.8	255	1	YWFD_BACSU
29	332	23.7	249	1	DHG2_STRVN
30	331.5	23.7	261	1	DHG_BACME
31	331.5	23.7	261	1	DHG_BACSU
32	330	23.6	247	1	YD50_MYCTU
33	329.5	23.6	264	1	DHG1_BACME

34	327.5	23.4	261	1	DHGA_BACME
35	326.5	23.3	255	1	HDHA_ECOLI
36	326.5	23.3	268	1	TRN1_DATST
37	326	23.3	244	1	FABG_BUCAI
38	324	23.2	245	1	NODG_RHIME
39	324	23.2	261	1	ACT3_STRCO
40	322.5	23.1	250	1	LIN3_PSEPA
41	322.5	23.1	261	1	DHG3_BACME
42	321.5	23.0	258	1	BDHA_RHIME
43	321	22.9	247	1	FABG_MYCTU
44	320.5	22.9	261	1	DHG4_BACME
45	319	22.8	244	1	FABG_VIBHA

ALIGNMENTS

RESULT 1					
HE27_HUMAN					
ID	HE27_HUMAN	STANDARD;	PRT;	257 AA.	
AC	Q13268;				
DT	01-NOV-1997 (Rel. 35, Created)				
DT	15-DEC-1998 (Rel. 37, Last sequence update)				
DT	20-AUG-2001 (Rel. 40, Last annotation update)				
DE	HEP27 PROTEIN (PROTEIN D).				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 1-18.				
RX	MEDLINE=96035881; PubMed=7556196;				
RA	Gabrielli F., Donadel G., Bensi G., Heguy A., Meili M.;				
RT	"A nuclear protein, synthesized in growth-arrested human				
RT	hepatoblastoma cells, is a novel member of the short-chain alcohol				
RT	dehydrogenase family.";				
RL	Eur. J. Biochem. 232:473-477(1995).				
RN	[2]				
RP	SEQUENCE OF 57-65; 118-123; 125-139 AND 175-182.				
RX	MEDLINE=9115312; PubMed=1847869;				
RA	Donadel G., Garzelli C., Frank R., Gabrielli F.;				
RT	"Identification of a novel nuclear protein synthesized in growth-				
RT	arrested human hepatoblastoma HepG2 cells.";				
RL	Eur. J. Biochem. 195:723-729(1991).				
CC	-1- FUNCTION: MAY INHIBIT CELL REPLICATION EITHER BY CATALYSING THE				
CC	OXIDATION OF ESTROGEN AND ANDROGEN OR BY CONVERTING CORTISONE IN				
CC	CORTISOL.				
CC	-1- SUBCELLULAR LOCATION: NUCLEAR.				
CC	-1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES				
CC	(SDR) FAMILY.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/				
CC	or send an email to license@isb-sib.ch).				
CC	-----				
DR	EMBL; U31875; AAA82048.1; ALT_INIT.				
DR	HSSP; P25529; 1AHH.				
DR	InterPro; IPR002198; ADH_short.				
DR	InterPro; IPR002347; Adh_short_C2.				
DR	Pfam; PF00106; adh_short; 1.				
DR	Pfam; PF00678; adh_short_C2; 1.				
DR	PRINTS; PR00081; GDRDH.				
DR	PROSITE; PS00061; ADH_SHORT; 1.				
KW	Oxidoreductase; Nuclear protein.				
FT	INIT_MET 0				
FT	NP_BIND 17 41				
FT	ACT_SITE 162 162				
FT	CONFLICT 57 57				
FT	CONFLICT 62 62				
FT	NAD OR NADP (BY SIMILARITY).				
FT	BY SIMILARITY.				
FT	L -> V (IN REF. 2).				
FT	L -> G (IN REF. 2).				

P10528	bacillus me
P35529	escherichia
P50165	datara stra
P57432	buchnera ap
P06234	rhizobium m
P16544	streptomyce
P30198	pseudomonas
P39484	bacillus me
O86034	rhizobium m
O48930	mycobacteri
P39485	bacillus me
P35336	vibrio harv

SQ SEQUENCE 257 AA; 27307; 612298E883725CAC CRC64;

Query Match 57.2%; Score 800.5; DB 1; Length 257;
Best Local Similarity 61.3%; Pred. No. 2.2e-52;
Matches 155; Conservative 47; Mismatches 50; Indels 1; Gaps 1;

QY 20 ASSGTRRPLANKVALVTASTDIGGFATARRLAQDGAHVVSRRKQNVQAVATLQGE 79
DB 1 STGIDRKGLANRAVAVVTGSGIGFATARRLAQDGAHVVSRRKQNVQAVATLQGE 60
QY 80 GLSVGTGCHVGAEDRERLAVATAVKLHGIGDILVSNAAVNPFFGIMDVTEVWDKTL 139
DB 61 GLSVAGIVCHVGAEDRERLAVATAVKLHGIGDILVSNAAVNPFFGIMDVTEVWDKTL 120
QY 140 INVKAPALMTKAVVPMERKGGVSVVSSIAAFSPGSPYVSKTALLGLTKTLATE 199
DB 121 VNVKSPALLSOLLPMENR-RGAVILVSSIAAANPVVALGVVYVSKTALLGLTKTLATE 179
QY 200 LAPRINRVNCLAPGLIKTSFRLMWDKKEESMKETLRIRLGEPEDCAGIVSFLCSED 259
DB 180 LAPKDIRVNCVVGIIKTSFRLMWDKKEESMKETLRIRLGEPEDCAGIVSFLCSED 239
QY 260 ASYITGETVVGG 272
DB 240 ASYVNGENAVAG 252

RESULT 2

FABG_BACSU STANDARD; PRT; 246 AA.
AC P51831.031733;
DT 01-OCT-1996 (Rel. 3.4) Created)
DT 15-DEC-1998 (Rel. 37; Last sequence update)
DT 20-AUG-2001 (Rel. 40; Last annotation update)
DE 3-OXOACYL-[ACYL-CARRIER PROTEIN] REDUCTASE (EC 1.1.1.100) (3-KETOACYL-
DE ACYL CARRIER PROTEIN REDUCTASE)
CN FABG.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=96326321; PubMed=8759840;
RA Morbidoni H.R., de Mendoza D., Cronan J.E. Jr.;
RT "Bacillus subtilis acyl carrier protein is encoded in a cluster of
RT lipid biosynthesis genes."
RL J. Bacteriol. 178:4794-4800(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RA Kunst F., Ogasawara N., Yoshikawa H., Danchin A.;
RN Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 1-172 FROM N.A.
RC STRAIN=168;
RX MEDLINE=98195738; PubMed=9534248;
RA Foulger D., Errington J.;
RT "A 28 kbp segment from the spovm region of the Bacillus subtilis 168
RT genome."
RL Microbiology 144:801-805(1998).
RN [4]
RP SEQUENCE OF 230-246 FROM N.A.
RC STRAIN=168;
RA Oguro A., Kakeshita H., Nakamura K., Yamane K.;
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: (3R)-3-HYDROXYACYL-[ACYL-CARRIER PROTEIN] +
CC NADP(+) -> 3-OXOACYL-[ACYL-CARRIER PROTEIN] + NADPH.
CC -1- PATHWAY: FIRST REDUCTION STEP IN THE FATTY ACID BIOSYNTHESIS
CC PATHWAY.
CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES

CC (SDR) FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: U59433; AAC44307.1; -
DR EMBL: Z99112; CAB13464.1; -
DR EMBL: Y13937; CAA74250.1; -
DR EMBL: D64116; BAA10974.1; -
DR HSSP: Q12634; 1YBV.
DR Subtilisin; BG11535; fabG.
DR InterPro: IPR002198; ADH_short.
DR InterPro: IPR002347; ADH_short_C2.
DR Pfam: PF00106; adh_short; 1.
DR Pfam: PF00678; adh_short_C2; 1.
DR PRINTS: PR00080; SDRFAMILY.
DR PRINTS: PR00081; GDRDH.
DR PROSITE: PS00061; ADH_SHORT; FALSE_NEG.
KW Fatty acid biosynthesis; Oxidoreductase; NADP; Complete proteome.
FT NP_BIND 9 33 NADP (BY SIMILARITY).
FT ACT_SITE 154 154 BY SIMILARITY.
FT CONFLICT 23 23 D -> A (IN REF. 1).
SQ SEQUENCE 246 AA; 26282 MW; C6A391167D3327DC CRC64;

Query Match 28.5%; Score 399; DB 1; Length 246;
Best Local Similarity 37.6%; Pred. No. 9.5e-23;
Matches 92; Conservative 50; Mismatches 99; Indels 4; Gaps 3;
QY 30 LANKVALVTASTDIGGFATARRLAQDGAHVVS-SRKQNVQAVATLQGEGLSVGTVC 88
DB 2 LNDKTAIVTGASRGISALDLAKSAGVNVYVSGNEAKANEVDEIKSMGRKAIVKA 61
QY 89 HVGAEDRERLAVATAVKLHGIGDILVSNAAVNPFFGIMDVTEVWDKTLDINVKAPALM 148
DB 62 DVSNPEDVQNMIKETLSVFSITDILVNNAGITR-DNLIMRKEDWDVDDVINLKGVC 120
QY 149 TKAVVPEMEKRGGSVIVSSIAAFSPGSPYVSKTALLGLTKTLAIELAPRINRVN 208
DB 121 TKAVTRQMKQSGRIINVSVGVSGNPGQANYAAKAGVIGLTKSKAKELASRNITVN 180
QY 209 CLAPGLIKTSFRLMWDKKEESMKETLRIRLGEPEDCAGIVSFLCSEDASYITGETV 268
DB 181 ATAPGFISTDMTDKL--AKVDQDEMLKQIPLARFCEPSDVSVVTFLLASEGARYVTGQTL 238
QY 269 VVGG 273
DB 239 HIDGG 243

RESULT 3

FABG_THEMA
ID FABG_THEMA STANDARD; PRT; 246 AA.
AC Q9X248;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE 3-OXOACYL-[ACYL-CARRIER PROTEIN] REDUCTASE (EC 1.1.1.100) (3-KETOACYL-
DE ACYL CARRIER PROTEIN REDUCTASE).
GN FABG OR TM1724.
OS Thermotoga maritima.
OC Bacteria; Thermotogales; Thermotoga.
OX NCBI_TaxID=2336;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MSB8 / DSM 3109;
RX MEDLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,

20-AUG-2001 (Rel. 40, Last annotation update)
 TROPINONE REDUCTASE-II (EC 1.1.1.236) (TR-II)
 TR2.
 Hyoscyamus niger (Henbane).
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asteridae; euasterids I; Solanales; Solanaceae; Hyoscyamus.
 NCBI_TaxID=4079;
 [1]
 SEQUENCE FROM N.A.
 TISSUE=Cultured root;
 RX MEDLINE=94120020; PubMed=8290643;
 RA Nakajima K., Hashimoto T., Yamada Y.;
 RT "cDNA encoding tropinone reductase-II from Hyoscyamus niger."
 RL Plant Physiol. 103:1465-1466(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Nakajima K., Hashimoto T.;
 RT "Tropinone reductase-II gene of Hyoscyamus niger";
 RL Submitted (Apr-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: CATALYZES THE STEREOSPECIFIC REDUCTION OF TROPINONE TO
 PSEUDOTROPINE.
 CC -1- CATALYTIC ACTIVITY: PSEUDOTROPINE + NADP(+) -> TROPINONE + NADPH.
 CC -1- PATHWAY: BIOSYNTHETIC PATHWAY OF TROPANE ALKALOIDS.
 CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
 (SDR) FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 or send an email to license@isb-sib.ch).
 CC
 CC EMBL; L20485; BAB09776.1;
 CC EMBL; AB026545; BAB85845.1;
 CC HSP; P47227; IBD.
 CC InterPro: IPR002198; ADH_short.
 CC InterPro: IPR002347; Adh_short_C2.
 CC Pfam; PF00106; adh_short; 1.
 CC Pfam; PF00678; adh_short_C2; 1.
 CC PRINTS; PR00081; GDHRDH.
 CC PROSITE; PS00061; ADH_SHORT; 1.
 CC Oxidoreductase; NADP.
 CC NP_BIND 13 37 NADP (BY SIMILARITY).
 CC ACT_SITE 159 159 BY SIMILARITY.
 CC SEQUENCE 260 AA; 28437 MW; 6CA7AF85CAA128FC CRC64;
 Query Match 27.1%; Score 379.5; DB:1; Length 260;
 Best Local Similarity 35.4%; Pred. No. 2.8e-21;
 Matches 90; Conservative 49; Mismatches 110; Indels 5; Gaps 3;
 QY 24 MTRDPLANKVALVASTDGIIGFAIRLAQDGAHVVSRRKQNVQAVATLOGEGLSV 83
 DB 1 MAGRWNEGCTALVTGSGRIGYGIYVELANLGASVYTCRSRQKDELCTQWRKGFV 60
 QY 84 TGTVCHVGKAEEDRLRVATAVK-LHGIDILVSNAAVPPFGSIMDVTVEWDTLDINV 142
 DB 61 EASVCDLSSRSEREFEKMTYNSPHGKLNILVNNAGI-VIYKEAKDTMEDYSHIMSNP 119
 QY 143 KAPALMTKAVVPEMEKRGGSVIVSSIAAFSPGSPYNSKALTGLTKTLAIELAP 202
 DB 120 EAYVHLSVAHPFLKASERGVNVYFISISGASALPYEAVYGATKGAMDOLTRCLAFEWAK 179
 QY 203 RNRVNCVLAFLGLTKTSRMLWMDKEESKMTL---RIRRGEPEDCAGIVSFLCSED 259
 DB 180 DNRVNGVGVGVATNSVEMTIDPEQKENLDKLDRCALRRMGEPKELAAVAVFLCPA 239
 QY 260 ASYTGETVVVGGG 273
 DB 240 ASYVTGGIIVDDGG 253

RESULT 6

Y325_THEME
 ID Y325_THEME STANDARD; PRT; 251 AA.
 AC Q9WYGO;
 DT 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE HYPOTHETICAL OXIDOREDUCTASE TM0325 (EC 1.1.1.236).
 GN TM0325.
 OS Thermotoga maritima.
 OC Bacteria; Thermotogales; Thermotoga.
 OX NCBI_TaxID=2336;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MSB8 / DSM 3109;
 RX MEDLINE=99287316; PubMed=10360571;
 RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
 RA Hart D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
 RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
 RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
 RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
 RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
 RT "Evidence for lateral gene transfer between Archaea and Bacteria from
 genome sequence of Thermotoga maritima."
 RL Nature 399:323-329(1999).
 CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
 (SDR) FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AE001714; AAD35412.1;
 CC TIGR; TM0325;
 CC InterPro: IPR002198; ADH_short.
 CC InterPro: IPR002347; Adh_short_C2.
 CC Pfam; PF00106; adh_short; 1.
 CC Pfam; PF00678; adh_short_C2; 1.
 CC PRINTS; PR00080; SDRFAMILY.
 CC PRINTS; PR00081; GDHRDH.
 CC PRINTS; PR01167; INSADHFAMILY.
 CC PROSITE; PS00061; ADH_SHORT; 1.
 CC Hypothetical protein; Oxidoreductase; Complete proteome.
 CC NP_BIND 10 34 NADP (BY SIMILARITY).
 CC ACT_SITE 152 152 BY SIMILARITY.
 CC SEQUENCE 251 AA; 26635 MW; 61C736A0F13564A5 CRC64;
 Query Match 26.7%; Score 373.5; DB:1; Length 251;
 Best Local Similarity 36.5%; Pred. No. 7.4e-21;
 Matches 91; Conservative 52; Mismatches 93; Indels 13; Gaps 5;
 QY 33 KVALVTASTDGIGFAIRRLAQAQGAHV---VSSRKQNVQAVATLOGEGLSVGTGVCH 89
 DB 6 KVLVTAGSGGIGKAAVMAFAERGAKVAINDISEKGETVELIKSMGGGAFFIGDV-- 63
 QY 90 VGRAEDRERLVAVAKLHGIDILVSNAAVPPFGSIMDVTVEWDTLDINVKAPALMT 149
 DB 64 ---AKDEQIVKKTVEFTGRLDILVNNAGIVP-YGNIETSEEDFDKTMVAVNKGFLLS 119
 QY 150 KAVVPEMEKRGGSVIVSSIAAFSPGSPYNSKALTGLTKTLAIELAPRNRVNC 209
 DB 120 KYAVEQMKGGGVIVNVSSSEAGLIGIPRCVYSVSRKALLGLTRSLAVDVIDVGIRVA 179
 QY 210 LAPGLIKTS--FSRMLWMDKEESKMTLRI---RIRRGEPEDCAGIVSFLCSEDASYITG 265
 DB 180 VCPGTTQSEGLMARVKASPNPELLKTKTSRIPMKRLGKBEEIAFALFAACDEAGFWTG 239


```

CC EMBL; L20474; AA33282.1;
DR PDB; 2AE1; 18-NOV-98.
DR PDB; 2AE2; 02-FEB-99.
DR InterPro; IPR002198; ADH_short.
DR InterPro; IPR002347; Adh_short_C2.
DR Pfam; PF00106; adh_short; 1.
DR Pfam; PF00678; adh_short_C2; 1.
DR PRINTS; PR00080; SDRFAMILY.
DR PRINTS; PR00081; GDRDH.
DR PROSITE; PS00061; ADH_SHORT; 1.
DR OXIDOREDUCTASE; NADP; 3D-structure.
KW NP_BIND 13 37 NADP (BY SIMILARITY).
FT ACT_SITE 159 159 BY SIMILARITY.
SQ SEQUENCE 260 AA; 28311 MW; 2DBF4963B2CCA303 CRC64;

Query Match 26.1%; Score 364.5; DB 1; Length 260;
Best Local Similarity 34.6%; Pred. No. 3.6e-20;
Matches 88; Conservative 50; Mismatches 111; Indels 5; Gaps 3;

Oy 24 MTRRDLANKVALVASTDGGIGFAIRLAQDGAHVYSSRRKQNVQAVATLQEGLSV 83
Db 1 WAGRNLEGCTALVTGSGRGIGYVEELASLGASVYTCRNQKELNDCLTQWRSGFKV 60

Oy 84 TGTCHVGKAEDEERLVAAT-AVKLHGDDILVSNAAVNPFFGSGIMDVTEYVMDKTLIN 142
Db 61 EASVCDLSRSRQELMTVANHPHGKLNILVNNAGI-VIYKEADYTVEDYSLIMSINF 119

Oy 143 KAPALMTKAVVPEMKRGSGSVYVSSAAFPSPGSPFPYNNVSKTALGLTKTLAIELAP 202
Db 120 EAAVHSLVLAHPFLKASRGNVVTFSSVSGALVYEAIVGATKGAMPDLTRCLAFENAK 179

Oy 203 RNIRVNCPLAGLTKTFSRMLWMDKEESMKEL--RIRLGPEDCAGIVSFLCSED 259
Db 180 DNRVNGVGPVATFSLVEMTIQDPEQENLNKLIDRCALRRMGPEKELAAWVAPLCPPA 239

Oy 260 ASYITGETVYVGGG 273
Db 240 ASYVTGQIIYVGGG 253

RESULT 9
UCPA_ECOLI STANDARD; PRO: 263 AA.
AC P37440; P76963; P77140;
DT 01-OCT-1994 (Rel. 30, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DE OXIDOREDUCTASE UCPA (EC 1.1.1.1)
GN UCPA OR B2426
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97426611; PubMed=9278503;
RA Blatner F.R., Plunkett, G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97349980; PubMed=9205837;
RA Yamamoto Y., Alba H., Baba T., Hayashi K., Inada T., Isono K.,
RA Itoh T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsuhashi N.,
RA Mizobuchi K., Mori H., Nakade S., Nakamura Y., Nashimoto H.,
RA Oshima T., Oiyama S., Saito N., Sampei G., Satoh Y., Sivasubram S.,

```

```

RA Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,
RA Yamagata S., Horiuchi T.;
RT "Construction of a contiguous 874-kb sequence of the Escherichia coli
RT -K12 genome corresponding to 50.0-68.8 min on the linkage map and
RT analysis of its sequence features.";
RL DNA Res. 4:91-113(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97385354; PubMed=9241368;
RA Sirko A., Weglenska A., Hryniewicz M.M., Hulanicka D.M.;
RT "Characterization of the Escherichia coli gene encoding a new member
RT of the short-chain dehydrogenase/reductase (SDR) family.";
RL Acta Biochim. Pol. 44:153-157(1997).
RN [4]
RP SEQUENCE OF 180-263 FROM N.A.
RC STRAIN=K12;
RX MEDLINE=90264335; PubMed=2189959;
RA Hryniewicz M.M., Sirko A., Palucha A., Boeck A., Hulanicka D.M.;
RT "Sulfate and thiosulfate transport in Escherichia coli K-12:
RT identification of a gene encoding a novel protein involved in
RT thiosulfate binding.";
RL J. Bacteriol. 172:3358-3366(1990).
RN [5]
RP IDENTIFICATION.
RX MEDLINE=95075659; PubMed=7984428;
RA Borodovsky M., Rudd K.E., Koonin E.V.;
RT "Intrinsic and extrinsic approaches for detecting genes in a
RT bacterial genome.";
RL Nucleic Acids Res. 22:4756-4767(1994).
CC -!- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
CC (SDR) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE000330; AAC75479.1; ALT_INIT.
CC EMBL; D90872; BAA16309.1; ALT_INIT.
CC EMBL; D90871; BAA16300.1;
CC EMBL; X99908; CAA68181.1;
CC EMBL; M32101; -. NOT_ANNOTATED_CDS.
CC HSSP; P14061; IFDV.
CC EcGene; EGI2133; ucpA.
CC InterPro; IPR002198; ADH_short.
CC InterPro; IPR002347; Adh_short_C2.
CC Pfam; PF00106; adh_short; 1.
CC Pfam; PF00678; adh_short_C2; 1.
CC PRINTS; PR00080; SDRFAMILY.
CC PRINTS; PR00081; GDRDH.
CC PROSITE; PS00061; ADH_SHORT; FALSE_NEG.
KW Oxidoreductase; Complete proteome.
FT CONFLICT 153 153 T -> Q (IN REF. 3).
SQ SEQUENCE 263 AA; 27850 MW; 2E201713357FDF41 CRC64;

Query Match 25.9%; Score 362.5; DB 1; Length 263;
Best Local Similarity 34.1%; Pred. No. 5.1e-20;
Matches 86; Conservative 53; Mismatches 104; Indels 9; Gaps 4;

Oy 30 LANKVALVASTDGGIGFAIRLAQDGAHVYSSRRKQNVQAVATLQEGLSVGTVCYH 89
Db 4 LTGXTALITGALQIGIGIARTFAHGHANLILDISPE-IEKLADELGRHRCATVAVD 62

Oy 90 VGKAEDEERLVAATVKLHGDDILVSNAAVNPFFGSGIMDVTEYVMDKTLINVKAPALMT 149
Db 63 VRDPASVAAAIKRAKEGRIDILVNNAGVCR-LGSFLQMSDDDDDFHIDINIKGWNVT 121

Oy 150 KAVPEMEKRGSGSVYVSSIAA-FSPSPGSPYNNVSKTALLGLTKTLAIELAPNIRVN 208

```

Db	76	ADPAQIRAYVAAAVQRYGTDVLVNNAG-RSGGATAEIADELWLWDVITTTNLTSVLMTK	134
Qy	151	AVVPE-----MEKRG-----GGSVVIVSSIAAFSPGFSPPYNSKTAALLGLTKTLAI	198
Db	135	EVLNAGGMLAKKRGIINIASTGGKQGVHIV-----PYSASKHGVVGLTKALGL	184
Qy	199	ELAPRNRVNCPLGLIKTGSFIRM-----LWMDKEKESMKETURI--RRLGEPEDCA	249
Db	185	ELARTGTTVAACPGFVETPMABRVREHYAGIQWVSEETFDRTINRPLUGRVVETREVA	244

RESULT	11
TRNL_DATST	
ID	TRNL_DATST
IR	STANDARD;
AC	P50162; PRT; 273 AA.
DT	01-OCT-1996 (Rel. 34, Created)
DT	01-OCT-1996 (Rel. 34, Last sequence update)
DT	20-AUG-2001 (Rel. 40, Last annotation update)
DE	TROPINONE REDUCTASE-I (EC 1.1.1.206) (TR-I) (TROPINE DEHYDROGENASE)
GN	TRL.
OS	Datura stramonium (Jimsonweed) (Common thornapple).
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC	Asteridae; euasterids I; Solanales; Solanaceae; Datura.
OX	NCBI_TaxID=4076;
RP	[1]
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.

CC -1- CATALYTIC ACTIVITY: TROPINE + NADP(+) = TROPINONE + NADPH.
CC
CC -1- PATHWAY: BIOSYNTHETIC PATHWAY OF TROPANE ALKALOIDS.
CC
CC -1- SUBUNIT: HOMODIMER.
CC
CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
CC (SDR) FAMILY.
CC
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial/
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
CC
CC ENBL; L20473; AAA33281.1; .
CC
CC PDB; 1AE1; 18-NOV-98.
CC
CC InterPro; IPR002198; ADH_short.
CC
CC InterPro; IPR002347; Adh_short_C2.
CC
CC Pfam; PF00106; adh_short; 1.
CC
CC Pfam; PF00678; adh_short_C2; 1.
CC
CC PRINTS; PR00081; GDHRDH.
CC
CC PROSITE; PS00061; ADH_SHORT; 1.
CC
CC Oxidoreductase; NADP; 3D-structure.
KW
CC NP_BIND 25 49 NADP (BY SIMILARITY).
CC FT ACT_SITE 171 171 BY SIMILARITY.
CC
CC

SQ SEQUENCE 273 AA; 29617 MW; 39A523EF04EA81F1 CRC64;

Query Match 25.5%; Score 357; DB 1; Length 273;
 Best Local Similarity 33.6%; Pred. No. 1.4e-19;
 Matches 89; Conservative 58; Mismatches 110; Indels 8; Gaps 4;

QY 15 NSVRMASSGMRDRPLANKVALVLTASTGIGFAIARRLAQDGAHVVSRRKQNVQDAVA 74
 DB 4 SKVSMNCCNCGRLSKGTTALTGSGGIGYVAIVEELAGLAGYVTCRSNEKELDECLE 63

QY 75 TLQGEGLSVGTGVCHVGAEDRERLVAATVAVKLHGG-IDILVSNAAVNPFFGSGIMDVTEEV 133
 DB 64 IWRKGLNVEGSCDLLSRTEDKLMQVAHVDFDCKNLILYNNAGV-VIHKEAKDFTKED 122

QY 134 WDKTLDINVKPALMTKAVVPEMEKRGSGVVIVSSIAAFSPGSPYNVSKTALLGLT 193
 DB 123 YNIIIMGTNEFAAYHLSQIAYPLLKASQNGNVLSSIAAGFSALPSVLSYASKAGINQMT 182

QY 194 KTLAELAPRINRVNCLAPG-----LITSTSRMLMDKKEESKMETLRIRRLGEPEDC 248
 DB 183 KSLACEAKDNIRVNSVAPGVIITLPTVETAIKKNPHQKEEIDNFIVKT-PMGRAGKQEV 241

QY 249 AGIVSFLCSEDASYITGETVVVGGG 273
 DB 242 SALIAFLCPAASYITGQIWDGG 266

RESULT 12

FABG_RICPR STANDARD; PRT; 241 AA.

ID AC P50941;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE 3-OXOACYL-[ACYL-CARRIER PROTEIN] REDUCTASE (EC 1.1.1.100) (3-KETOACYL-
 DE ACYL CARRIER PROTEIN REDUCTASE)
 GN FABG OR RP762.
 OS Rickettsia prowazekii.
 OC Bacteria; Proteobacteria; alpha subdivision; Rickettsiales;
 OC Rickettsiaceae; Rickettsiae; Rickettsia.
 OX NCBI_TaxID=782;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MADRID E;
 RX MEDLINE=99039499; PubMed=9823893;
 RA Andersson S.G.E., Zomorodipour A., Andersson J.O.,
 RA Sicheritz-Ponten T., Alsmark U.C.M., Podowski R.M., Naeslund A.K.,
 RA Eriksson A.-S., Winkler H.H., Kurland C.G.;
 RT "The genome sequence of Rickettsia prowazekii and the origin of
 RL mitochondria."
 RL Nature 396:133-140(1998).
 RN [2]
 RP SEQUENCE OF 1-41 FROM N.A.
 RC STRAIN=MADRID E;
 RX MEDLINE=94179117; PubMed=8132476;
 RA Dunkin S.M., Winkler H.H., Wood D.O.;
 RT "Isolation and characterization of the Rickettsia prowazekii reca
 RL gene."
 RL J. Bacteriol. 176:1777-1781(1994).
 RN [3]
 RP IDENTIFICATION.
 RX MEDLINE=96259307; PubMed=8662004;
 RA Andersson S.G.E., Sharp P.M.;
 RT "Codon usage and base composition in Rickettsia prowazekii."
 RL J. Mol. Evol. 42:525-536(1996).
 CC -1- CATALYTIC ACTIVITY: (3R)-3-HYDROXYACYL-[ACYL-CARRIER PROTEIN] +
 CC NADP(+) = 3-OXOACYL-[ACYL-CARRIER PROTEIN] + NADPH
 CC -1- PATHWAY: FIRST REDUCTION STEP IN THE FATTY ACID BIOSYNTHESIS
 CC PATHWAY.
 CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
 CC (SDR) FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)
 CC or send an email to license@isb-sib.ch)

EMBL; AJ235273; CAAL15190.1;
 DR EMBL; U01959; ; NOT_ANNOTATED_CDS.
 DR HSP; Q12634; 1YBV.
 DR InterPro; IPR002198; ADH_short.
 DR InterPro; IPR002347; Adh_short_C2.
 DR Pfam; PF00106; adh_short; 1.
 DR PRINTS; PR00080; SDRFAMILY.
 DR PRINTS; PR00081; GDHRDH.
 DR PRINTS; PR01167; INSADHFAMILY.
 DR PROSITE; PS00061; ADH_SHORT; 1.
 KW Fatty acid biosynthesis; Oxidoreductase; NADP; Complete proteome.
 FT NP_BIND 11 35 NADP (BY SIMILARITY).
 FT ACT_SITE 148 148 BY SIMILARITY.
 SQ SEQUENCE 241 AA; 25759 MW; E43B8711545B8295 CRC64;

Query Match 25.3%; Score 354; DB 1; Length 241;
 Best Local Similarity 33.6%; Pred. No. 1.9e-19;
 Matches 82; Conservative 61; Mismatches 91; Indels 10; Gaps 4;

QY 30 LANKVALVASTDGIGFAIARRLAQDGAHVVSRRKQNVQDAVATLOGEGLSVTGVCH 89
 DB 4 LTGKTSLTIGASSGIGSAIARLLHKLKSGVITSGNEEKLKSLGNALKD---NYTIEVCN 60

QY 90 VKAEDRERLVAATVAVKLHGGIDILVSNAAVNPFFGSGIMDVTEEVWDKLDINVKAPALMT 149
 DB 61 LANKKEECNLISKT---SNLDILVCNAGITSDDTAIR-MKDQDFKVIDINKANFILN 115

QY 150 KAVPEMEKRGSGVVIVSSIAAFSPGSPYNVSKTALLGLTTLAELAPRINRVNC 209
 DB 116 REAKKMKLOKRYGRIINISSIVGAGNPGQANYCASKAGLIGMTKLSYEVAITGVNA 175

QY 210 LAPGLIKTSFRMLMDKKEESKMETLRIRRLGEPEDCAGIVSFLCSEDASYITGETVV 269
 DB 176 VAPGFIKSDMTDKL--NEKQREAIQKIPGLTYGIPEDVAVAVAFSLASNNASYITQTLH 233

QY 270 VGGG 273
 DB 234 VNGG 237

RESULT 13

Y4LA_RHISN STANDARD; PRT; 278 AA.

ID AC P55541;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE PUTATIVE SHORT-CHAIN TYPE DEHYDROGENASE/REDUCTASE Y4LA (BC 1.-.-.-).
 GN Y4LA.
 OS Rhizobium sp. (strain NGR234).
 OG Plasmid sym PNGR234a.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Rhizobium.
 OX NCBI_TaxID=394;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97305956; PubMed=9163424;
 RA Freiberg C.A., Fellay R., Bairoch A., Broughton W.J., Rosenthal A.,
 RA Perret X.;
 RT "Molecular basis of symbiosis between Rhizobium and legumes."
 RL Nature 387:394-401(1997).
 CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
 CC (SDR) FAMILY.

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; AE000082; AAB91754.1;
 CC HSP; P29132; 1DFT
 CC InterPro; IPR002198; ADH_short.
 CC InterPro; IPR002347; Adh_short_C2.
 CC Pfam; PF00106; adh_short; 1.
 CC Pfam; PF00678; adh_short_C2; 1.
 CC PRINTS; PR00080; SDRFAMILY.
 CC PRINTS; PR00081; GDHRDH.
 CC PROSITE; PS00061; ADH_SHORT; 1.
 KW Hypothetical protein; Oxidoreductase; Plasmid.
 FT NP_BIND 10 34 NAD OR NADP (BY SIMILARITY).
 FT ACT_SITE 159 159 BY SIMILARITY.
 SQ SEQUENCE 278 AA; 28743 MW; 1D0105625BE9DF2B CRC64;

Query Match 25.3%; Score 354; DB 1; Length 278;
 Best Local Similarity 35.5%; Pred. No. 2.3e-19;
 Matches 87; Conservative 46; Mismatches 106; Indels 6; Gaps 3;

QY 35 KVALVTASTDGTGIGFAIARLAODGAHVYSSRRQQNVQAVATLOGEGLSVTGTVCHVGK 92
 DB 7 KVVAVTGAGAGIGKACALAIAREGGRVVVADLDGSAIACTAQIAAEAGNALAMADIAD 66
 QY 93 AEDRELAVATVAVKLHGGIDILYSNAA---VNPFFGSIMDVTEVWDKTLIDINVKAPALMT 149
 DB 67 AQVAALEFETAEHRFGVGDVLLVNNASAMHLPRDRAILDLDLVAVDQMTATNLRLGLCC 126
 QY 150 KAVVPEMKRGSGSVIVSSIAAFSPGSPYVNVSKTALLGLTKTLATLAPRNRVNC 209
 DB 127 RQAIPRMIARGGAIVNMSSCOGLSDTAQTSYAVSKAAMNLSASLATQYGHAIQRCNA 186
 QY 210 LAPGLIKTSFRLWMDKEESMKETLRIRRLGPEDCAGIVSFLCSEDASYITGETV 268
 DB 187 VAPGLIMT--ERLLAKDKCMRHLSRHOLLPRVGHPEVVAALVALLSDDASFTIGQVV 244
 QY 269 VVGGG 273
 DB 245 CIDGG 249

RESULT 14
 Y019_THEME STANDARD; PRT; 256 AA.
 AC Q56318;
 DT 01-NOV-1997 (Rel. 35; Created)
 DT 30-MAY-2000 (Rel. 39; Last sequence update)
 DT 20-AUG-2001 (Rel. 40; Last annotation update)
 DE PUTATIVE OXIDOREDUCTASE.TM0019 (EC 1.1.1.17).
 GN TM0019.
 OS Thermotoga maritima.
 CC Bacteria; Thermotogales; Thermotoga.
 OX NCBI_TaxID=2336;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MSB8 / DSM 3109;
 RX MEDLINE=96125254; PubMed=8550425;
 RA Kletzin A., Adams M.
 RT "Molecular and phylogenetic characterization of pyruvate and 2-
 RT ketoisovalerate ferredoxin oxidoreductases from Pyrococcus furiosus
 RT and pyruvate ferredoxin oxidoreductase from Thermotoga maritima";
 RL J. Bacteriol. 178:248-257 (1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MSB8 / DSM 3109;

RX MEDLINE=99287316; PubMed=10360571;
 RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
 RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
 RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
 RA Stewart A.M., Sutton M.D., Pratt M.S., Phillips C.A., Richardson D.,
 RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
 RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
 RT "Evidence for lateral gene transfer between Archaea and Bacteria from
 RT genome sequence of Thermotoga maritima";
 RL Nature 399:323-329 (1999).
 CC -!- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
 CC (SDR) FAMILY.
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).

CC EMBL; X85171; CAAS9459.1;
 CC EMBL; AE001690; AAD35113.1;
 CC HSP; P19932; 2HSD.
 CC TIGR; TM0019;
 CC InterPro; IPR002198; ADH_short.
 CC InterPro; IPR002347; Adh_short_C2.
 CC Pfam; PF00106; adh_short; 1.
 CC Pfam; PF00678; adh_short_C2; 1.
 CC PRINTS; PR00080; SDRFAMILY.
 CC PRINTS; PR00081; GDHRDH.
 CC PRINTS; PRO1167; INSADHFAMILY.
 CC PROSITE; PS00061; ADH_SHORT; 1.
 KW Hypothetical protein; Oxidoreductase; NADP; Complete proteome.
 FT NP_BIND 9 33 NADP (BY SIMILARITY).
 FT ACT_SITE 153 153 BY SIMILARITY.
 FT CONFLICT 130 136 RGGVII -> TWRSDH (IN REF. 1).
 SQ SEQUENCE 256 AA; 28078 MW; D68160BD17980C6B CRC64;

Query Match 24.9%; Score 349; DB 1; Length 256;
 Best Local Similarity 35.2%; Pred. No. 4.9e-19;
 Matches 89; Conservative 45; Mismatches 103; Indels 16; Gaps 5;

QY 30 LANKVALVTASTDGTGIGFAIARLAODGAHVYSSRRQQNVQAVATLOGEGLSVTGTVCH 89.
 DB 2 LEGKVVAVTGGGGGIGAAIAQLFAENGKMWIAEIDEEAGVEREEMLRERGLDVTGVKTD 61
 QY 90 VGKAEDRELAVATVAVKLHGGIDILYSNAAVNPFFGSIMDVTEVWDKTLIDINVKAPALMT 149
 DB 62 VADENSVKMMVKRTVEIYGGVDVLVNNAAVMS-VKSIFERPLEENERVIRVNLTPGYICS 120
 QY 150 KAVVPEMKRGSGSVIVSSIAAFSPGSPYVNVSKTALLGLTKTLATLAPRNRVNC 209
 DB 121 RYCAEMIKRGGVVINTIASTRAFSEPTPEYSASKGGLVALTHSLVSLRYHVRVVS 180
 QY 210 LAPGLIKTSFRLWMDKEESMKETLR-----IRRLGPEDCAGIVSFLC-SEDA 260
 DB 181 ISPGWIETS---EW-KKSLRKKKPDRLPIDHEQHPAGRVGNPLDLIAHLCLVFLADDEKA 234
 QY 261 SYITGETVVVGGG 273
 DB 235 GFITGTNFTVDDG 247

RESULT 15
 DHG2_BACSU STANDARD; PRT; 258 AA.
 ID DHG2_BACSU
 AC P80869;
 DT 01-NOV-1997 (Rel. 35; Created)
 DT 15-JUL-1999 (Rel. 38; Last sequence update)
 DT 20-AUG-2001 (Rel. 40; Last annotation update)
 DE GLUCOSE 1-DEHYDROGENASE II (EC 1.1.1.47) (GLCDH-II) (GENERAL

```

Db      185 NAIAPGTIATESN-----VDTKKEESROKQLKIPMKAFGKPPEEVAAMLVSEASYVT   240
Qy      265 GETVVVGGGT-----PSRL 278
        |||:|||||
Db      241 GATLFVDGGMTLFPSQL 257

```

Search completed: January 22, 2002, 15:22:36
Job time: 191 sec

[illegible]

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 22, 2002, 15:19:10 ; Search time 23.8 Seconds
(without alignments)
1708.559 Million cell updates/sec

Title: US-09-866-034-2
Perfect score: 1399
Sequence: 1 MHKAGLLGLCARAWNSVRMA.....DASYITGETVVVGGGTPSRL 278

Scoring table: BLOSOM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues
Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_17.*
1: sp.archaea.*
2: sp.bacteria.*
3: sp.fungi.*
4: sp.human.*
5: sp.invertebrate.*
6: sp.mammal.*
7: sp.mhc.*
8: sp.organelle.*
9: sp.phase.*
10: sp.plant.*
11: sp.rodent.*
12: sp.virus.*
13: sp.vertebrate.*
14: sp.unclassified.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1399	100.0	278	4 Q9BT22	Q9bt22 homo sapien
2	1395	99.7	278	4 Q9NV08	Q9nv08 homo sapien
3	1301	93.0	260	4 Q9H3N5	Q9h3n5 homo sapien
4	1293	92.4	260	4 Q9S162	Q9s162 homo sapien
5	1123	80.3	260	6 Q9GKX2	Q9gkx2 oryctolagus
6	1086	77.6	260	14 Q9EQU4	Q9equ4 mus musculus
7	1086	77.6	260	11 Q99LE2	Q99lb2 mus musculus
8	824.5	58.9	280	4 Q9H2R2	Q9h2r2 homo sapien
9	787	56.3	282	11 Q9D3W7	Q9d3w7 mus musculus
10	716	51.2	216	11 Q9D2B3	Q9d2u3 mus musculus
11	664	47.5	317	5 Q9VRJ4	Q9vrj4 drosophila
12	636	45.5	257	5 Q16619	Q16619 caenorhabdi
13	614	43.9	260	5 Q93790	Q93790 caenorhabdi
14	588.5	42.1	254	10 Q9S9W2	Q9s9w2 arabidopsis
15	468	33.5	341	11 Q9DC11	Q9dcl1 mus musculus
16	459.5	32.8	260	2 Q93308	Q93308 mycobacteri
17	454	32.5	254	2 Q9A3X5	Q9a3x5 caulobacter
18	409	29.2	255	2 Q9I2R7	Q9i2r7 pseudomonas
19	389.5	27.8	253	2 Q9K3Y8	Q9k3y8 streptomyce

20	388	27.7	246	2 Q9KA03	Q9ka03 bacillus ha
21	384.5	27.5	273	2 Q9F5J1	Q9f5j1 streptomyce
22	381.5	27.3	266	2 Q9EW14	Q9ew14 streptomyce
23	381.5	27.3	282	1 Q9S564	Q9s564 pyrococcus
24	381	27.2	255	1 Q9HQ41	Q9hq41 halobacteri
25	377	26.9	264	10 Q9LHT0	Q9lht0 arabidopsis
26	376	26.9	261	2 Q54812	Q54812 streptomyce
27	374	26.7	243	2 Q9CHF7	Q9chf7 lactococcus
28	371	26.5	260	2 Q9A8R5	Q9a8r5 caulobacter
29	371	26.5	263	1 Q9V212	Q9v212 pyrococcus
30	370	26.4	249	2 Q9KEB5	Q9keb5 bacillus ha
31	369.5	26.4	246	2 Q9A7P5	Q9a7p5 caulobacter
32	367	26.2	240	2 Q9X429	Q9x429 streptococc
33	366	26.2	261	2 Q54280	Q54280 saccharopol
34	365.5	26.1	248	2 Q31680	Q31680 bacillus su
35	364	26.0	254	2 Q9RT26	Q9rt26 deinococcus
36	363	25.9	274	10 Q96457	Q96457 hyocycamus
37	360.5	25.8	261	2 Q54176	Q54176 streptomyce
38	359	25.7	248	2 Q9KJF1	Q9kjf1 thauera aro
39	358.5	25.6	256	2 P95273	P95273 mycobacteri
40	357.5	25.6	259	2 Q9S3U5	Q9s3u5 bacteroides
41	357.5	25.6	268	1 Q9HLN6	Q9hln6 thermoplas
42	357	25.5	248	2 Q9JXR1	Q9jxr1 neisseria m
43	357	25.5	248	2 Q9JW61	Q9jw61 neisseria m
44	356	25.4	264	10 Q9AR59	Q9ar59 solanum tub
45	354.5	25.3	262	2 Q9L9F8	Q9l9f8 streptomyce

ALIGNMENTS

RESULT 1
Q9BT22 PRELIMINARY; PRT; 278 AA.
AC Q9BT22
DT 01-JUN-2001 (TREMREL. 17, Created)
DT 01-JUN-2001 (TREMREL. 17, Last sequence update)
DT 01-JUN-2001 (TREMREL. 17, Last annotation update)
DE PEROXISOMAL SHORT-CHAIN ALCOHOL DEHYDROGENASE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG CARCINOMA;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC003019; AAH03019.1;
SQ SEQUENCE 278 AA; 29537 MW; 3B06A229E1BBE47B CRC64;

Query Match	100.0%;	Score	1399;	DB	4;	Length	278;
Best Local Similarity	100.0%;	Pred. No.	1.8e-96;				
Matches	278;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
Qy	1	MHKAGLLGLCARAWNSVRMASSGTMTRDPLANKVALVTASTDGIGFAIARRLAQDGAHV	60				
Db	1	MHKAGLLGLCARAWNSVRMASSGTMTRDPLANKVALVTASTDGIGFAIARRLAQDGAHV	60				
Qy	61	VSSRKQNVQDAVATLQGEGLSVGTGVCVHGKAEDRRLVATKVLHGIDILVSNAAVN	120				
Db	61	VSSRKQNVQDAVATLQGEGLSVGTGVCVHGKAEDRRLVATKVLHGIDILVSNAAVN	120				
Qy	121	PFPGSIMDYTEEVWTKTLDINVKAPALMTKAVVPKEMKGGGWSVIVSSIAAFSPSPGFS	180				
Db	121	PFPGSIMDYTEEVWTKTLDINVKAPALMTKAVVPKEMKGGGWSVIVSSIAAFSPSPGFS	180				
Qy	181	PYNVSTALLGLTKTIAELAPRINVINCLAPGLIKTFSRMLWMDKEESKKTILRIR	240				
Db	181	PYNVSTALLGLTKTIAELAPRINVINCLAPGLIKTFSRMLWMDKEESKKTILRIR	240				
Qy	241	RLGPEPCAGIVSFLCSEDASYITGETVVVGGGTPSRL	278				

```
|||||
Db 241 RLGPEDCAGIVSLCSEDASYITGETVVVGGTSPRL 278

RESULT 2
Q9NV08 PRELIMINARY; PRT: 278 AA.
AC Q9NV08;
DT 01-OCT-2000 (TrEMBLrel.15, Created)
DT 01-OCT-2000 (TrEMBLrel.15, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel.17, Last annotation update)
DE CDNA FLJ11008 FIS, CLONE PLACE1003100, MODERATELY SIMILAR TO HEP27
DE PROTEIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,
RA Tanase T., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K.,
RA Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,
RA Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y., Oshima A.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES (SDR)
CC FAMILY.
DR EMBL; AK001870; BAA91953.1; -.
DR InterPro: IPR002198; Adh_short.
DR InterPro: IPR002347; Adh_short_C2.
DR Pfam; PF00106; adh_short; 1.
DR Pfam; PF00678; adh_short_C2; 1.
DR PRINTS; PR00080; SDRFAMILY.
DR PROSITE; PS00061; ADH_SHORT; UNKNOWN_1.
KW Oxidoreductase.
SQ SEQUENCE 278 AA; 29509 MW; 6B0096BBA89152A0 CRC64;

Query Match 99.78%; Score 1395; DB 4; Length 278;
Best Local Similarity 99.6%; Pred. No. 3.6e-96;
Matches 277; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MHKAGLLGLCARAVNSVRMSGTRRDPLANKVALVTASTDGIGFAIARRLAQDGAHV 60
DB 1 MHKAGLLGLCARAVNSVRMSGTRRDPLANKVALATASTDGIGFAIARRLAQDGAHV 60
QY 61 VSRKQNVQAVATLQEGLSVTGTVCHVGAEDRERLAVATAVKLHGIDILVSNAAVN 120
DB 61 VSRKQNVQAVATLQEGLSVTGTVCHVGAEDRERLAVATAVKLHGIDILVSNAAVN 120
QY 121 PFGSMDVTEVWDKTLIDINVKAPALMTKAVVPEMEKRGSGSVIVSSIAAFSPGFS 180
DB 121 PFGSMDVTEVWDKTLIDINVKAPALMTKAVVPEMEKRGSGSVIVSSIAAFSPGFS 180
QY 181 PYNYSKTALLGLTKTLATEAPRNVNCLAPGLIKTSFRMLWMDKEESMKETLRIR 240
DB 181 PYNYSKTALLGLTKTLATEAPRNVNCLAPGLIKTSFRMLWMDKEESMKETLRIR 240
QY 241 RLGPEDCAGIVSLCSEDASYITGETVVVGGTSPRL 278
DB 241 RLGPEDCAGIVSLCSEDASYITGETVVVGGTSPRL 278

RESULT 3
Q9H3N5 PRELIMINARY; PRT: 260 AA.
AC Q9H3N5;
DT 01-MAR-2001 (TrEMBLrel.16, Created)
DT 01-MAR-2001 (TrEMBLrel.16, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel.17, Last annotation update)
DE NADPH-DEPENDENT RETINOL DEHYDROGENASE/REDUCTASE.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA Furukawa A., Ohnishi T., Huang D., Araki N., Ichikawa Y.;
RA "cDNA cloning and characterization of peroxisomal short-chain
RT dehydrogenase / reductase that reduce all-trans retinal to retinol.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES (SDR)
CC FAMILY.
DR EMBL; AB045131; BAB18775.1; -.
DR InterPro: IPR002198; Adh_short.
DR InterPro: IPR002347; Adh_short_C2.
DR Pfam; PF00106; adh_short; 1.
DR Pfam; PF00678; adh_short_C2; 1.
DR PRINTS; PR00080; SDRFAMILY.
DR PROSITE; PS00061; ADH_SHORT; UNKNOWN_1.
KW Oxidoreductase.
SQ SEQUENCE 260 AA; 27572 MW; 0BD2E0C0D2E37D08 CRC64;

Query Match 93.08%; Score 1301; DB 4; Length 260;
Best Local Similarity 100.0%; Pred. No. 3.2e-89;
Matches 260; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 MASSGMRTRDPLANKVALVTASTDGIGFAIARRLAQDGAHVVSRRKQNVQAVATLQ 78
DB 1 MASSGMRTRDPLANKVALVTASTDGIGFAIARRLAQDGAHVVSRRKQNVQAVATLQ 60
QY 79 EGLSVTGTCHVGAEDRERLAVATAVKLHGIDILVSNAAVNPFSGIMDVTEVWDKTL 138
DB 61 EGLSVTGTCHVGAEDRERLAVATAVKLHGIDILVSNAAVNPFSGIMDVTEVWDKTL 120
QY 139 DINVKAPALMTKAVVPEMEKRGSGSVIVSSIAAFSPGFSVYNVSKTALLGLTKTLAI 198
DB 121 DINVKAPALMTKAVVPEMEKRGSGSVIVSSIAAFSPGFSVYNVSKTALLGLTKTLAI 180
QY 199 ELAPRNVNCLAPGLIKTSFRMLWMDKEESMKETLRIRRLGEPEDCAGIVSLCSE 258
DB 181 ELAPRNVNCLAPGLIKTSFRMLWMDKEESMKETLRIRRLGEPEDCAGIVSLCSE 240
QY 259 DASVITGETVVVGGTSPRL 278
DB 241 DASVITGETVVVGGTSPRL 260

RESULT 4
Q95162 PRELIMINARY; PRT: 260 AA.
AC Q95162;
DT 01-MAY-1999 (TrEMBLrel.10, Created)
DT 01-MAY-1999 (TrEMBLrel.10, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel.17, Last annotation update)
DE PEROXISOMAL SHORT-CHAIN ALCOHOL DEHYDROGENASE.
GN SCAD-SRL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Fransen M., Van Veldhoven P.P., Subramani S.;
RT "Identification of two novel mammalian proteins with a C-terminal
RT peroxisomal targeting signal by using the pVI-phase display
RT technology.";
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES (SDR)
CC FAMILY.
DR EMBL; AF044127; AAD02292.1; -.

```

DR HSSP: P50163; 2AEL1
 DR InterPro; IPR002198; Adh_short.
 DR InterPro; IPR002347; Adh_short_C2.
 DR Pfam; PF00106; adh_short_C1.
 DR Pfam; PF00106; adh_short_C2; 1.
 DR PRINTS; PR00080; SDRFAMILY.
 DR PROSITE; PS00061; ADH_SHORT; UNKNOWN_1.
 KW Oxidoreductase.
 SQ SEQUENCE 260 AA; 27602 MW; 83687CACD82B9B5 CRC64;

Query Match 92.4%; Score 1293; DB 4; Length 260;
 Best Local Similarity 99.2%; Pred. No. 1.3e-88;
 Matches 258; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 19 MASSGTRDRPLANKVALVTAATDGI GFAIARRLAODGAHVSVSSRKQONVDQAVATLQ 78
 Db 1 MASSGTRDRPLANKVALVTAATDGI GFAIARRLAODGAHVSVSSRKQONVDQAVATLQ 60
 Qy 79 EGLSVGTGVCHVGAEDRERLAVATKVLHGGIDILVSNAAVNPFGSIMDVTVEEVDKTL 138
 Db 61 EGLSVGTGVCHVGAEDRERLAVATKVLHGGIDILVSNAAVNPFGSIMDVTVEEVDKTL 120
 Qy 139 DINVKAPALMTKAVVPEMEKRGGSVIVSSIAAFSPSPGSPYVNSKTALLGLTKTLAI 198
 Db 121 DINVKAPALMTKAVVPEMEKRGGSVIVSSIAAFSPSPGSPYVNSKTALLGLTKTLAI 180
 Qy 199 ELAPNRVNCPLAGLIKTSFRLMWDKEESMKETLRIRRLGEPEDCAGIVSFLCSE 258
 Db 181 ELAPNRVNCPLAGLIKTSFRLMWDKEESMKETLRIRRLGEPEDCAGIVSFLCSE 240
 Qy 259 DASYYTGTGVVGGTSPSL 278
 Db 241 DASYYTGTGVVGGTSPSL 260

RESULT 5
 Q9GKX2 ID Q9GKX2 PRELIMINARY; PRT; 260 AA.
 AC Q9GKX2;
 DT 01-MAR-2001 (TREMREL. 16, Created)
 DT 01-MAR-2001 (TREMREL. 16, Last sequence update)
 DT 01-JUN-2001 (TREMREL. 17, Last annotation update)
 DE NADPH-DEPENDENT RETINOL DEHYDROGENASE/REDUCTASE.
 GN RABNROR.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER;
 RA Furukawa A., Ohnishi T., Huang D., Araki N., Ichikawa Y.;
 RT "cDNA cloning and characterization of peroxisomal short-chain
 RT dehydrogenase / reductase that reduce all-trans retinal to retinol.";
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES (SDR)
 CC FAMILY.
 DR EMBL; AB045133; BAB18776.1;
 DR InterPro; IPR002198; Adh_short.
 DR InterPro; IPR002347; Adh_short_C2.
 DR Pfam; PF00106; adh_short_C1.
 DR Pfam; PF00678; adh_short_C2; 1.
 DR PRINTS; PR00080; SDRFAMILY.
 DR PROSITE; PS00061; ADH_SHORT; UNKNOWN_1.
 KW Oxidoreductase.
 SQ SEQUENCE 260 AA; 27429 MW; 5B0585B58911B90C CRC64;

Query Match 80.3%; Score 1123; DB 6; Length 260;
 Best Local Similarity 85.4%; Pred. No. 5.6e-76;
 Matches 222; Conservative 18; Mismatches 20; Indels 0; Gaps 0;

Qy 19 MASSGTRDRPLANKVALVTAATDGI GFAIARRLAODGAHVSVSSRKQONVDQAVATLQ 78
 Db 1 MASSGTRDRPLANKVALVTAATDGI GFAIARRLAODGAHVSVSSRKQONVDQAVATLQ 60
 Qy 79 EGLSVGTGVCHVGAEDRERLAVATKVLHGGIDILVSNAAVNPFGSIMDVTVEEVDKTL 138
 Db 61 EGLSVGTGVCHVGAEDRERLAVATKVLHGGIDILVSNAAVNPFGSIMDVTVEEVDKTL 120
 Qy 139 DINVKAPALMTKAVVPEMEKRGGSVIVSSIAAFSPSPGSPYVNSKTALLGLTKTLAI 198
 Db 121 DINVKAPALMTKAVVPEMEKRGGSVIVSSIAAFSPSPGSPYVNSKTALLGLTKTLAI 180
 Qy 199 ELAPNRVNCPLAGLIKTSFRLMWDKEESMKETLRIRRLGEPEDCAGIVSFLCSE 258
 Db 181 ELAPNRVNCPLAGLIKTSFRLMWDKEESMKETLRIRRLGEPEDCAGIVSFLCSE 240
 Qy 259 DASYYTGTGVVGGTSPSL 278
 Db 241 DASYYTGTGVVGGTSPSL 260

RESULT 6
 Q9EQU4 ID Q9EQU4 PRELIMINARY; PRT; 260 AA.
 AC Q9EQU4;
 DT 01-MAR-2001 (TREMREL. 16, Created)
 DT 01-MAR-2001 (TREMREL. 16, Last sequence update)
 DT 01-JUN-2001 (TREMREL. 17, Last annotation update)
 DE NADPH-DEPENDENT RETINOL DEHYDROGENASE/REDUCTASE.
 GN MOUNROR.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57 BLAK; TISSUE=LIVER;
 RA Furukawa A., Ohnishi T., Huang D., Araki N., Ichikawa Y.;
 RT "cDNA cloning and characterization of peroxisomal short-chain
 RT dehydrogenase / reductase that reduce all-trans retinal to retinol.";
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES (SDR)
 CC FAMILY.
 DR EMBL; AB045133; BAB18776.1;
 DR InterPro; IPR002198; Adh_short.
 DR InterPro; IPR002347; Adh_short_C2.
 DR InterPro; IPR003015; HLH_Myc.
 DR Pfam; PF00106; adh_short; 1.
 DR Pfam; PF00678; adh_short_C2; 1.
 DR PRINTS; PR00080; SDRFAMILY.
 DR PROSITE; PS00038; HELIX_LOOP_HELIX; UNKNOWN_1.
 KW Oxidoreductase.
 SQ SEQUENCE 260 AA; 27726 MW; 005C5F19AC539CF6 CRC64;

Query Match 77.6%; Score 1086; DB 11; Length 260;
 Best Local Similarity 81.5%; Pred. No. 3.1e-73;
 Matches 212; Conservative 22; Mismatches 26; Indels 0; Gaps 0;

Qy 19 MASSGTRDRPLANKVALVTAATDGI GFAIARRLAODGAHVSVSSRKQONVDQAVATLQ 78
 Db 1 MASSGTRDRPLANKVALVTAATDGI GFAIARRLAODGAHVSVSSRKQONVDQAVATLQ 60
 Qy 79 EGLSVGTGVCHVGAEDRERLAVATKVLHGGIDILVSNAAVNPFGSIMDVTVEEVDKTL 138
 Db 61 EGLSVGTGVCHVGAEDRERLAVATKVLHGGIDILVSNAAVNPFGSIMDVTVEEVDKTL 120
 Qy 139 DINVKAPALMTKAVVPEMEKRGGSVIVSSIAAFSPSPGSPYVNSKTALLGLTKTLAI 198
 Db 121 DINVKAPALMTKAVVPEMEKRGGSVIVSSIAAFSPSPGSPYVNSKTALLGLTKTLAI 180
 Qy 199 ELAPNRVNCPLAGLIKTSFRLMWDKEESMKETLRIRRLGEPEDCAGIVSFLCSE 258
 Db 181 ELAPNRVNCPLAGLIKTSFRLMWDKEESMKETLRIRRLGEPEDCAGIVSFLCSE 240

```

Db 181 ELAPKNIRVNCLAPGLIKTRFSSVLWEKAREDFIKAMQIRRLGKPEDCAGIVSFLCSE 240
QY 259 DASVITGETVVGSGTPSRL 278
Db 241 DASVINGETVVGSGTPSRL 260
RESULT 7
Q99LB2 PRELIMINARY; PRT; 260 AA.
AC Q99LB2;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE SIMILAR TO PEROXISOMAL SHORT-CHAIN ALCOHOL DEHYDROGENASE.
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=MAMMARY TUMOR;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC003484; AA03484.1;
SQ SEQUENCE 260 AA; 27754 MW; 021733159D2BC3C9 CRC64;

Query Match 77.6%; Score 1086; DB 11; Length 260;
Best Local Similarity 81.5%; Pred. No. 3.1e-73;
Matches 212; Conservative 22; Mismatches 26; Indels 0; Gaps 0;

QY 19 MASSGTRDRPLANKVALVTASTDGIGFAIARRLAODGAHVVSRRKQONVQAVATLOG 78
Db 1 MASSGLTRNPLSNKVAHVASTDGIGFAIARRLAEDGAHVVSRRKQONVQAVATLOG 60
QY 79 EGLSVTGTVCHVGKADRELRVATVAKLHGIDILVSNAAVNPFGSIMDVTEEVWDKTL 138
Db 61 EGLSVTGTVCHVGKADRELRVATVAKLHGIDILVSNAAVNPFGSIMDVTEEVWDKTL 120
QY 139 DINVKAPALMTKAVVPEMEKRGSGSVIVSSIAAFSPSPGFPYNSKATLGLTKTLAI 198
Db 121 SINVTATAMIKAVVPEMEKRGSGSVIVSGVAGTRFSLGPNVNSKATLGLTKFNAA 180
QY 199 ELAPNIRVNCLAPGLIKTSRMLWMDKEESMKETLIRRLGPEPCAGIVSFLCSE 258
Db 181 ELAPKNIRVNCLAPGLIKTRFSSVLWEKAREDFIKAMQIRRLGKPEDCAGIVSFLCSE 240
QY 259 DASVITGETVVGSGTPSRL 278
Db 241 DASVINGETVVGSGTPSRL 260

RESULT 8
Q9H2R2 PRELIMINARY; PRT; 280 AA.
AC Q9H2R2;
DT 01-MAR-2001 (TReMBLrel. 16, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE SHORT-CHAIN ALCOHOL DEHYDROGENASE.
GN HEP27.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Pellegrini S., Censini S., Guidotti S., Covacci A., Gabrielli F.;
RT "Human Hep27 chromosomal gene";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES (SDR)
CC FAMILY.

```

```

DR EMBL; AF244132; AAC33703.1;
DR InterPro; IPR002198; ADH_short.
DR InterPro; IPR002347; Adh_short_C2.
DR Pfam; PF00106; adh_short_1.
DR Pfam; PF00678; adh_short_C2; 1.
DR PRINTS; PR00080; SDRFAMILY.
DR PROSITE; PS00061; ADH_SHORT; UNKNOWN_1.
KW Oxidoreductase.
SQ SEQUENCE 280 AA; 29926 MW; FE76ED9CB28AB95C CRC64;

Query Match 58.9%; Score 824.5; DB 4; Length 280;
Best Local Similarity 61.6%; Pred. No. 9.9e-54;
Matches 162; Conservative 47; Mismatches 51; Indels 3; Gaps 2;

QY 10 CARANVSVMASSTGTRDRPLANKVALVTASTDGIGFAIARRLAODGAHVVSRRKQONV 69
Db 16 CARL--SVRMSSTGIDRGKGLANRVAVVGTSGTIGFAIARRLAODGAHVVSRRKQONV 73
QY 70 DOAVATLOGEGSLVGTGCHVGKADRELRVATVAKLHGIDILVSNAAVNPFGSIMDV 129
Db 74 DRAMAKLOGEGSLVAGIVCHVGKADRELRVAKALEHCGGVDFLVCAGYNPLVGTSLGT 133
QY 130 TEEVWDKTLIDINVKAPALMTKAVVPEMEKRGSGSVIVSSIAAFSPSPGFPYNSKATL 189
Db 134 SEQINDKILSVNVKSPALLSOLLPPYMNENR- RGAIVLVSSIAAYNPVVALGVNVSKTAL 192
QY 190 LGITKTALTEIAPNRINRVNCLAPGLIKTSRMLWMDKEESMKETLIRRLGPEPCDCA 249
Db 193 LGLTRTALTEIAPNRINRVNCLAPGLIKTSRMLWMDKEESMKETLIRRLGPEPCDCA 252
QY 250 GIVSFLCSDASVITGETVVGSG 272.
Db 253 GIVSFLCSDASVINGENIATVAG 275

RESULT 9
Q9D3M7 PRELIMINARY; PRT; 282 AA.
AC Q9D3M7;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)
DE 5430405K24RIK PROTEIN.
GN 5430405K24RIK.
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=HEAD;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yananaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staabli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection";
RL Nature 409:685-690(2001).

```



```

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Matti B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasmann D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RA "The genome sequence of Drosophila melanogaster."
RT Science 287:2185-2195(2000).
CC -!- SIMILARITY: TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES (SDR)
CC FAMILY
DR EMBL; AF003567; AAF50801.1; -.
DR HSP; P50162; IAE1.
DR FlyBase; FBgn0035588; CG10672.
DR InterPro; IPR002198; Adh_short.
DR InterPro; IPR002347; Adh_short_C2.
DR Pfam; PF00106; adh_short; 1.
DR Pfam; PF00678; adh_short_C2; 1.
DR PRINTS; PR00080; SDRFAMILY.
DR PROSITE; PS00061; ADH_SHORT; 1.
DR Oxidoreductase.
KW OXIDOREDUCTASE.
SQ SEQUENCE 317 AA; 33617 MW; 0D8C30A7E64EE4E3 CRC64;

Query Match 47.5%; Score 664; DB 5; Length 317;
Best Local Similarity 50.9%; Pred. No. 1e-41;
Matches 136; Conservative 50; Mismatches 79; Indels 2; Gaps 1;

QY 12 RAWNSVRMASCWTRRPLANKVALVASTDGTGIGFAIRRAQDGAHVVSRRKQNVQD 71
Db 53 RLSSSSOSTAGTKMR--LAGKVAVVASTDGTGIGFAIRRAQDGAHVVSRRKQNVDS 110
QY 72 AVATLOGEGLSVTCVGHVGAEDRERLIVATVAKLHGIDILVSNAAVNPFFGSIIMDYTE 131
Db 111 ALAELRLNVLNVHGLKCHVSEPDQRKQLFEETISKFKGLNLTLSNAAVNPFFGSIIMDYTE 170
QY 132 EWDKTLIDINVKAPALMKAIVPEMEKRGSGSVIVSVISIAAFSPGFSYVNSKTLALIG 191
Db 171 KWMDKIFDVNVKSSYLAKAELPLRQKNSSIVFVSSIAGYDAFELLGAYSVSKTALIG 230
QY 192 LTKTLALAPRNRVNCAPGLTKTSFRNMLMDKEESMKETLIRIRLGEPEDCAGI. 251
Db 231 LTRAAAKDLAPEGIRVNCAPGVIRTKFSKALYENESANEALSKIPMGRIGTSEENAGV 290
QY 252 VSFCLSEDASYITGETVVGCGTSPRL-278
Db 291 VSFVLVEDAGYITGESVAGGGMTARL 317

RESULT 12
ID Q16619 PRELIMINARY; PRT; 257 AA.
AC Q16619;
DT 01-JAN-1998 (T-EMBLrel. 05, Created)
DT 01-JAN-1998 (T-EMBLrel. 05, Last sequence update)
DT 01-JUN-2001 (T-EMBLrel. 17, Last annotation update)
DE F36H9.3 PROTEIN.
GN F36H9.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RC SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;

```

```

RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Alnsough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Consey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA Lightnings J., Lloyd C., McMurray A., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,
RA Smalldon N., Smith A., Sonhammer E., Staden R., Sulston J.,
RA Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA Watson A., Weinstein L., Wilkinson-Sproat J., Wohlman P.,
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans."
RN Nature 368:32-38(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RC Dante M., Kramer J., Twyman B.;
RL Submitted (AUG-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RC Waterston R.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES (SDR)
CC FAMILY
DR EMBL; AF016668; AAB66092.1; -.
DR HSP; P50162; IAE1.
DR InterPro; IPR002198; Adh_short.
DR InterPro; IPR002347; Adh_short_C2.
DR Pfam; PF00106; adh_short; 1.
DR Pfam; PF00678; adh_short_C2; 1.
DR Oxidoreductase.
KW OXIDOREDUCTASE.
SQ SEQUENCE 257 AA; 27617 MW; FBF35DE9DEBEE1F CRC64;

Query Match 45.5%; Score 636; DB 5; Length 257;
Best Local Similarity 50.6%; Pred. No. 9.3e-40;
Matches 126; Conservative 49; Mismatches 74; Indels 0; Gaps 0;

QY 30 LANKVALVASTDGTGIGFAIRRAQDGAHVVSRRKQNVQDVAATLOGEGLSVGTVCH 89
Db 9 LTRVALVASTGTGIGFAIRRAQDGAHVVSRRKQNVQDVAATLOGEGLSVGTVCH 68
QY 90 VGKAEDRERLIVATVAKLHGIDILVSNAAVNPFFGSIIMDYTEVWDKTLIDINVKAPALMT 149
Db 69 VGNKSDTRKTLIDFTDRFTKDLIDLVSNAAVNPFFGSIIMDYTEVWDKTLIDINVKAPALMT 128
QY 150 KAVVPEMEKRGSGSVIVSVISIAAFSPGFSYVNSKTLALIGLTKLALAPRNRVNC 209
Db 129 KEAVPHELESGRNVVSVSSVAGYSPMNEICAYSVKTKTLGLSKSLALNARNIRVNS 188
QY 210 LAPGLIKTSFRNMLMDKEESMKETLIRIRLGEPEDCAGIVSFLCSEDASYITGETV 269
Db 189 IAPGLIOTDFESQVLFSDSEKQKWLSQLAQRFRGDPDECAEAVAFVLSDEASYISGETIG 248
QY 270 VGGTSPRL 278
Db 249 INGMHARI 257

RESULT 13
Q93790 PRELIMINARY; PRT; 260 AA.
ID Q93790;
AC Q93790;
DT 01-FEB-1997 (T-EMBLrel. 02, Created)
DT 01-MAY-1999 (T-EMBLrel. 10, Last sequence update)
DT 01-JUN-2001 (T-EMBLrel. 17, Last annotation update)
DE F54F3.4 PROTEIN.
GN F54F3.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.

```


OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Percy C.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES (SDR)
CC FAMILY
DR EMBL; 279696; CAB01974.1; JOINED.
DR EMBL; 281592; CAB01974.1; JOINED.
DR EMBL; 281592; CAB04734.1; JOINED.
DR EMBL; 279696; CAB04734.1; JOINED.
DR HSP; P50163; 2AE1.
DR InterPro: IPR002198; ADH_short.
DR InterPro: IPR002347; ADH_short_C2.
DR Pfam: PF00106; adh_short; 1.
DR Pfam: PF00678; adh_short_C2; 1.
DR PRINTS; PR00080; SDRFAMILY.
KW Oxidoreductase.
SQ SEQUENCE 260 AA; 27590 MW; E43FD36F5EDBA7F4 CRC64;

Query Match 43.9%; Score 614; DB 5; Length 260;
Best Local Similarity 48.0%; Pred. No. 4.1e-38;
Matches 120; Conservative 56; Mismatches 70; Indels 4; Gaps 2;

QY 33 KVALVTASTDGIGFAIARLAODGAHVVSRRKQNVDOAVATLQEGLS-VTGTVCHVG 91
Db 11 KVAIVTAATKGIGLAERLLDEGASVWIGSRQKNVDENIEYLNKGLTKVAGIAGHTA 70
QY 92 KAEDRLRLVATVKGIGIDILVSNAAVNPFGSIMDVTVEWVKDITLDINVKAPALMTKA 151
Db 71 STDQKKLVDFTLQKFGKINILNNGHINPAFGHILEVSDQVWDLKLFVNVKRGAFQMTKL 130
QY 152 VYPEMEKRGSGVVIYSSIAAFSPSPGFPYNNVSKTALLGLTKTLAELAPRINRVNCL 211
Db 131 VPHIAKEGGGAILFNAVSAYKSPPGIAYGVTKTLVGLTRALAMGLAKDNIRVNGIA 190
QY 212 PGLIKTSFRLMLW---MDKEESMKETLRIRLGEPEDCAGIVSFLCSEDASYITGETV 268
Db 191 PGVIKTKMSQVLWDGDEDAEKELTDQETALGLRGVDPDCAGTAVLAYLSDSSYITGEM 250
QY 269 VVGGTGPSRL 278
Db 251 ILAGGVQARL 260

RESULT 14
Q9S9W2 ID Q9S9W2 PRELIMINARY; PRT; 254 AA.
AC Q9S9W2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE T1J24.9 PROTEIN (AT4G05530 PROTEIN).
GN T1J24.9 OR AT4G05530.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Ali J., Bauer C., Nguyen C., Ducks G.;
RT "The sequence of A. thaliana T1J24.";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Washu;
RT "The A. thaliana Genome Sequencing Project.";
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
RN [3]

RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Waterston R.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Lamar B., Stoneking T., Stumpf J., Mewes H.W., Lemcke K.,
RA Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES (SDR)
CC FAMILY
DR EMBL; AF147263; AAD48959.1; -;
DR EMBL; AL161503; CAB81095.1; -;
DR HSP; P29132; IDFI.
DR InterPro: IPR002198; ADH_short.
DR InterPro: IPR002347; ADH_short_C2.
DR Pfam: PF00106; adh_short; 1.
DR Pfam: PF00678; adh_short_C2; 1.
DR PRINTS; PR00080; SDRFAMILY.
DR PROSITE; PS00061; ADH_SHORT; UNKNOWN_1.
KW Oxidoreductase.
SQ SEQUENCE 254 AA; 26765 MW; 54297E4D1D095372 CRC64;

Query Match 42.1%; Score 588.5; DB 10; Length 254;
Best Local Similarity 50.6%; Pred. No. 3.1e-36;
Matches 126; Conservative 35; Mismatches 85; Indels 3; Gaps 2;

QY 30 LANKVALVTASTDGIGFAIARLAODGAHVVSRRKQNVDOAVATLQEGLSVTGTVCH 89
Db 9 LEKVAIVTASTGIGIGTEREGLEGASVVSRRKQNVDOAVAKLKGIDAYGLVCH 68
QY 90 VKAEDRLRLVATVKGIGIDILVSNAAVNPFGSIMDVTVEWVKDITLDINVKAPALMT 149
Db 69 VSNQHRRLVKTQYKYGKIDIVVCNAAANPSTDPISSKEAVLDKLWEINVKSSILL 128
QY 150 KAVVPEMEKRGSGVVIYSSIAAFSPSPGFPYNNVSKTALLGLTKTLAELAPRINRVN 209
Db 129 QDMAPHEK--GSSVIFITSAGFSPQAMAYGVTKTALLGLTKALAEAMAP-DTRVNA 185
QY 210 LAPGLIKTSFRLMLWMDKEESMKETLRIRLGEPEDCAGIVSFLCSEDASYITGETV 269
Db 186 VAFGVPTHFASFTGSSEVREGIEETLLNRLTGTDMAAAAFASDDSSYITGETLV 245
QY 270 VGGGTGPSRL 278
Db 246 VAGGMPSRL 254

RESULT 15
Q9DCL1 ID Q9DCL1 PRELIMINARY; PRT; 141 AA.
AC Q9DCL1;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE DNA SEGMENT, CHR 14, UNIVERSITY OF CALIFORNIA AT LOS ANGELES 2.
GN D14UCLA2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=KIDNEY;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Harai A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,

Search completed: January 22, 2002, 15:22:21
Job time: 191 sec